

南 華 大 學

資訊管理學系

碩士論文

應用資料採礦技術探討國小學生學習背景  
對學業成績的影響

Applying Data Mining to Investigate the Influence of the  
Background on the Academic Performance of Elementary  
School Children.



研 究 生：韓岱芳

指 導 教 授：邱宏彬 博士

中 華 民 國 103 年 6 月 22 日

# 南 華 大 學

## 資 訊 管 理 學 系 碩 士 學 位 論 文

應用資料採礦技術探討國小學生學習背景  
對學業成績的影響

研究生：韓岱芳

經考試合格特此證明

口試委員：陳世諤

邱宏彬

李翔諤

指導教授：邱宏彬

系主任(所長)：洪銘建

口試日期：中華民國 103 年 6 月 22 日

南華大學碩士班研究生  
論文指導教授推薦函

資訊管理系碩士班韓岱芳君所提之論文  
應用資料採礦技術探討國小學生學習背景對學  
業成績的影響

係由本人指導撰述，同意提付審查。

指導教授

邱宏村  
103年6月 日

## 南華大學資訊管理學系碩士論文著作財產權同意書

立書人：\_\_\_\_\_韓岱芳\_\_\_\_\_之碩士畢業論文

中文題目：應用資料採礦技術探討國小學生學習背景對學業成績的影響

英文題目：Applying Data Mining to Investigate the Influence of the Background on the Academic Performance of Elementary School Children.

指導教授：邱宏彬 博士

學生與指導老師就本篇論文內容及資料其著作財產權歸屬如下：

- 共同享有著作權
- 共同享有著作權，學生願「拋棄」著作財產權
- 學生獨自享有著作財產權

學生：\_\_\_\_\_韓岱芳\_\_\_\_\_ (請親自簽名)

指導老師：\_\_\_\_\_邱宏彬\_\_\_\_\_ (請親自簽名)

中華民國 103 年 6 月 22 日

## 誌 謝

本文能夠順利完成，首先要感謝指導教授邱宏彬老師，在論文寫作期間的鼓勵與指導，我才能克服許多研究上的困難與盲點，老師的教誨和體諒包容，學生永銘在心。同時，也感謝口試委員李翔詣教授及陳張宗榮教授對本文提供寶貴的經驗與建議，讓論文能更趨於完備，特此致謝。

感謝親愛的老婆明怡，在這段時間擔負起照顧孩子的責任，讓我能有較多的時間專心的完成論文；也感謝親愛的家人，體諒我這兩年常無法參與家庭活動的心情。

另外，也要感謝資管所的教授們，在二年碩專班的課程學習中，以其專業的學術涵養，帶我們遨遊各學術領域的天空，分享人生的經驗。

回首兩年碩專班的求學歷程，過程辛苦卻很充實。其中最寶貴資產就是認識資管所的這一群良師益友，謝謝大家！

# 應用資料採礦技術探討國小學生學習背景對 學業成績的影響

學生：韓岱芳

指導教授：邱宏彬 博士

南 華 大 學 資 訊 管 理 學 系 碩 士 班

## 摘 要

本研究旨在探討國小學生的個人學習背景因素與學業成績之相關性。使用資料採礦之關聯規則找出國小四、五、六年級學生之學期成績與個人學習背景因素之關係。

資料來源以雲林縣某國小 102 學年度四、五、六年級學生共 113 人為樣本，資料取自學生個人學習背景因素調查結果及學務管理系統。研究方法採用 Microsoft SQL Server 2012 資料採礦之關聯法則來進行資料分析。

研究結果顯示成績好的學生有以下規則：每日平均學習時間為 1 小時以上；平均每天使用電腦時間在 1 小時以內；平均一星期裡面有 3 天以上能主動學習；平均一星期裡面有 3 天以上作息是有規律的；學習遇到困難時，勇於發問與虛心求助的平均機率需高於五成；參加補習家教的科目數在 2 科上；認真學習的科目在 3 科以上。

關鍵詞：學業成績、學習背景因素、資料採礦、關聯規則。

# Applying Data Mining to Investigate the Influence of the background on the Academic Performance of Elementary School Children.

Student : Han,Tai-Fang

Advisor : Dr. Chiu,Hung-Pin

Department of Information Management  
The Graduated Program  
Nan-Hua University

## ABSTRACT

The aim of this study is to explore the correlation between elementary school student's background and their academic achievement. Use association rule technique in data mining to analyze the data. The result is for reference and assisting students to make the planning of study.

The researcher employed grades four, five, six students composed of 113 people 102 Academic Year at an elementary school located in Yunlin County. The data was retrieved from findings of student's background and the Student Affairs Management System. The research methods used the association rule of Microsoft SQL Server 2012 to proceed with the data analysis process.

The study shows that students of good achievements have the following characteristics. It is more than 1 hours on average every day that they study .It is less than 1 hours on average every day that they use computer. It is more than 3 days on average every week that they study voluntarily and regularly. When having difficulties in studying, they often ask for help. It is more than 2 subjects that they participate in extra-curricular tutoring. It is more than 3 subjects that they study seriously.

Keywords : Academic Performance , Student's Background ,

Data Mining, Association Rule.

# 目 錄

誌謝.....	i
中文摘要.....	ii
英文摘要.....	iii
目錄.....	iv
表目錄.....	v
圖目錄.....	vi
第一章 緒論.....	1
第一節 研究動機.....	1
第二節 研究目的.....	2
第三節 研究範圍與限制.....	2
第四節 研究步驟.....	3
第五節 論文架構.....	4
第二章 文獻探討.....	6
第一節 影響學生學業成就之因素.....	6
第二節 資料採礦技術.....	9
第三節 關聯規則.....	10
第三章 研究方法.....	15
第一節 資料來源與資料篩選.....	15
第二節 資料的預處理.....	19
第三節 資料採礦工具 SQL Server 2012 簡介.....	21
第四節 演算法參數設定.....	22
第四章 結果與討論.....	24
第一節 資料特性分析.....	24
第二節 關聯法則模型建置過程.....	26
第三節 資料採礦-關聯規則的彙整與分析.....	39
第五章 結論.....	48
第一節 研究結論.....	48
第二節 後續研究建議.....	49
參考文獻.....	50
附錄一.....	52



# 表 目 錄

表 2-1	影響學業成就因素相關研究.....	8
表 2-2	早餐店的豆漿與三明治銷售關係表.....	12
表 2-3	交易資料表.....	13
表 3-1	成績轉換等第對照表.....	16
表 3-2	學生學期成績表.....	16
表 3-3	學生的個人學習背景因素之問卷與選項.....	17
表 3-4	學生個人學習背景因素變項值調查結果.....	19
表 3-5	學生學期成績主表.....	19
表 3-6	學生個人學習背景因素資料表.....	20
表 3-7	系統的演算法參數表.....	22
表 3-8	學生各成績等第的人數分配表.....	23
表 4-1	學期成績等第人數分佈表.....	25
表 4-2	「成績等第=優」參數設定表.....	40
表 4-3	「成績等第=優」規則彙整表.....	40
表 4-4	「成績等第=甲」參數設定表.....	41
表 4-5	「成績等第=甲」規則彙整表.....	42
表 4-6	「成績等第=乙」規則彙整表.....	43
表 4-7	「成績等第=丙」規則彙整表.....	44
表 4-8	成績等第與個人學習背景因素關聯規則彙整表.....	45

## 圖 目 錄

圖 2-1	資料表掃描與過濾結果(一).....	14
圖 2-2	資料表掃描與過濾結果(二).....	14
圖 4-1	學生性別比例分佈圖.....	24
圖 4-2	新增資料庫 ScoreDMdb.....	26
圖 4-3	匯入 Scorefactortb.xlsx 資料表.....	27
圖 4-4	設定 scorefactor\$ 資料表適當的資料類型.....	28
圖 4-5	設定主鍵畫面一.....	29
圖 4-6	設定主鍵畫面二.....	29
圖 4-7	新增採礦專案 ScoreDMProj.....	30
圖 4-8	ScoreDMProj 專案新增資料來源.....	31
圖 4-9	ScoreDMProj 專案連接 ScoreDMdb 資料庫.....	31
圖 4-10	ScoreDMProj 專案新增資料來源檢視.....	32
圖 4-11	ScoreDMProj 專案連接 scoremain\$ 和 scorefactor\$資料 表.....	33
圖 4-12	scorefactor\$ 資料表新增具名計算.....	34
圖 4-13	新增採礦結構.....	35
圖 4-14	選取 Microsoft 關聯規則技術.....	35
圖 4-15	指定 scoremain\$和 scorefactor\$資料表類型.....	36
圖 4-16	指定分析使用的資料行.....	37
圖 4-17	指定分析資料行的資料類型為「類別變數」.....	38
圖 4-18	設定關聯規則演算法參數.....	39
圖 4-19	「成績等第=優」相依性網路圖.....	41
圖 4-20	「成績等第=甲」相依性網路圖.....	42

圖 4-21 「成績等第=乙」相依性網路圖.....43  
圖 4-22 「成績等第=丙」相依性網路圖.....44



# 第一章 緒論

本論文的緒論共分成五小節，第一節旨在說明研究之動機；第二節則是敘述研究之目的；第三節是界定研究的範圍及限制；第四節則是提出本論文之研究步驟與流程；最後，第五節則是本篇論文之架構。

## 第一節 研究動機

學生的學業成就高低，代表學生客觀而重要的學習結果（鄭燕祥，2001），是學生在校學習表現的具體指標之一，向來為家長、教師及學校所重視。依據教育心理學所言，與學生個人學習成就最有直接關係的因素，便是個人投入學習的直接行為和心理特質（余民寧，2006），吳天泰（1996）指出影響兒童學業成就最主要的因素是心理特質與環境，心理特質類同於學生個人因素；在同樣學習情境下，有的學生能專注於課業或其它活動的學習，獲得學習成就，有的學生面對學習，卻感到困難、挫折、缺乏自信心，無法提起興趣，以致學習動力不足、學習態度不佳，造成學業成就低落。因此，探討影響學生學業成就的個人層面因素，將有助於釐清其學習的歷程，做為其他學生自我學習的借鏡。

『多花時間讀書才會有好成績！』是家長與老師常向學生說的一句話。可見學生讀書時間是影響學業成績的重要因素，「學習時間」將是本論文探討的主要因素之一。依據國教司統計 93 年基本學力測驗滿分的學生學習狀況後發現：智力表現並不是影響學習成就的重要因素；平均每天看電視的時間在 1 小時以內約占七成；平均每天使用電腦時

間在 1 小時以內約占七成(李義平，民 96)。爰此，本論文將探討學生個人的學習背景因素對其學期成績之影響，並運用資料採礦技術(Data Mining)找出學生個人的學習背景因素與學生學期成績之間的關係，進而提供學生學習方式的建議與協助學生做好學習的規畫。

## 第二節 研究目的

本研究採用資料採礦工具裡的關聯法則 (Association Rule) 演算法來找出學生個人的學習背景因素與學期成績之間的關係，一來因為研究資料的筆數不多；二來資料內容很少增減異動，所以用關聯法則裡最常用的 Apriori 演算法來做資料分析是最簡潔有效的方法。

本論文預期所要達成的目的如下：

- 壹、 運用資料採礦技術找出學生個人的學習背景因素與學期成績之間的關係，以作為教師、學生、家長參考。
- 貳、 提供學生課後學習活動及時間安排上的建議，以協助其做好適當的學習規劃。

## 第三節 研究範圍與限制

### 壹、 研究範圍

本論文以雲林縣某國民小學之四、五、六年級普通班學生為研究對象。研究範圍以研究對象的 102 學年度第一學期國語、英語、數學、自然與生活科技、社會等五科學期成績的平均與學生個人學習背景因素的關係。

### 貳、 學生的學習背景因素之範圍

本論文探討學生的學習背景因素之範圍包含 1 學習時間、2 休

閒時間、3 使用電腦時間、4 下課時間的利用、5 學習動機(主動學習)、6 學習計畫與執行力、7 作息規律性、8 虛心求助頻率、9 校外補習科目數以及 10 學習態度(認真學習)等十個變項。

#### 參、 研究限制

本論文研究對象僅限於雲林縣某個案國小，囿於學校背景因素不同，研究結果概括性小，無法過度類推至其他的學校之學生。

### 第四節 研究步驟

本論文以雲林縣某國民小學之四、五、六年級普通班學生 102 學年度第一學期國語、英語、數學、自然與生活科技、社會等五科學期成績的平均，作為研究使用之資料庫，以資料採礦中之關聯法則為採礦工具，來討論學生個人的學習背景因素與學期成績之間的關係，並依循下列步驟來完成本研究。研究步驟包含：研究動機及目的確認、文獻探討、擬定研究方法及架構、資料採礦與分析、結論與建議，其說明如下：

- 壹、 研究動機及目的確認：確認本研究之動機與目的，以確認資料蒐集之方向及研究之進行與操作。
- 貳、 資料蒐集與文獻回顧：蒐集相關文獻以瞭解研究之相關知識並尋求適合本研究之技術及工具。
- 參、 擬定研究方法及架構：針對所研究之問題及所運用之工具，擬定適合之研究方法及架構。
- 肆、 分析資料的取得：本研究以雲林縣某國民小學之四、五、六年級普通班學生 102 學年度第一學期國語、英語、數學、自

然與生活科技、社會等五科學期成績，作為研究使用之資料庫，由該校學務管理系統取得學期成績原始資料，並設計線上問卷取得學生學習背景因素之變項值。

伍、資料前置處理：將學生成績資料做適當的預處理，並排除未填問卷等遺漏資料後，得到學生學期成績資料與個人學習背景因素資料做為資料採礦資料庫。

陸、資料採礦分析：運用關聯法則為資料採礦工具，進行相關資料之分析，找出學生個人學習背景因素與學期成績之關聯規則。

柒、結論：針對關聯法則分析所得到的規則進行整理，推敲、解釋背後可能的原因，並提出學生學習方式的建議以協助學生做好學習的規畫。

## 第五節 論文架構

本篇論文共分為五章，各章節內容之安排概略說明如下：

### 第一章緒論

逐步說明本研究之研究動機、研究目的、研究範圍與限制、研究步驟以及本論文架構。

### 第二章文獻探討

分別說明影響學生學業成就之因素與相關研究、資料採礦技術與相關應用、關聯規則原理與 Apriori 演算法。

### 第三章研究方法

針對資料來源與資料篩選、資料預處理、資料採礦工具 SQL

SERVER 2012 以及演算法參數設定逐一做概略介紹。

#### 第四章結果與討論

針對資料做敘述性統計，整理出各成績等第人（次）數分佈表，並說明進行關聯規則模型建置過程以及資料採礦-關聯規則的分析結果。

#### 第五章結論

針對關聯法則分析所得到的規則進行整理，推敲、解釋背後可能的原因，最後提出結論與建議。





## 第二章 文獻探討

本章主要探討與本論文相關之理論與研究，全章共分為三節，第一節說明影響學生學業成就之因素與相關研究；第二節介紹資料採礦技術與相關應用；第三節說明關聯規則原理與 Apriori 演算法。

### 第一節 影響學生學業成就之因素

#### 壹、學業成就的意義

余民寧（2006）則針對「學習成就」進行明確定義，其認為就廣義的定義來說，為各種在校期間學生的學習記錄資料，而狹義的定義則是各學科的學期成績或各學科綜合後的平均學期成績。依此，本論文所指學業成就，即以學生在校學業成績為主，係指學生在校的各項課業學習，所獲致學習的結果，其範圍包括國語、英語、數學、自然與生活科技和社會等領域的學業評量成績，亦即學生該五項領域的學期學業總成績。

#### 貳、影響學業成就相關因素之理論

余民寧（2006）提出影響學習成就因素的統整性理論模型，以引用潛在變項模式的路徑關係，企圖解釋和概括過去局部的理論模型，其指出影響學習成就的來源，可分成直接效果與間接效果，其中造成學習成就最直接的效果，可歸納出五大類：學生個人背景因素、學生家庭背景因素、教師教學背景因素、學校管理背景因素及政府教育政策因素等。

Finn(1972)提出「期望網」理論，其指出家長對子女教育的關心與期望，會影響子女的自我期望，而同儕關係、學校教師與他人的期望亦會促進個人的自我期望，並產生交互影響，增強自我概念，進而影響其學習成就(引自周新富，1999)。

#### 參、影響學業成就之因素

國內外對於研究學業成就之相關文獻非常豐富，Gill 與 Spilka (1962) 將影響學業成就的層面，歸納為四大因素，分別為 1. 智力。2. 學生人格。3. 父母對子女學業的態度。4. 社會或經濟地位(引自王財印,2000);而國內學者王文科(1991)認為影響學業成就的因素，可區分為內在因素：指個人的智力、動機、人格特質等，及外在因素：指教學方法、學習團體、教師等；郭生玉(1981)則提出除了智力因素外，影響學業成就的包括心理因素：個人動機、人格適應、態度及學習習慣；生理因素：視覺、聽覺及一般健康狀態；社會因素：家庭背景、教育態度、家長職業及社區文化價值；教育因素：教學方法、課程與教材等。吳天泰(1996)則認為影響學業成就最主要的個人因素：是心理特質與環境，心理特質類同於前述個人因素。陳怡華(2001)、李美慧(2004)及郭春悅(2007)等人亦認為影響學業成就的因素可以歸納為個人因素和環境因素二個面向。茲將影響學業成就因素相關研究整理如表 2-1。

表 2-1 影響學業成就因素相關研究

研究者	影響原因
Gill 與 Spilka (1962)	<ol style="list-style-type: none"> <li>1. 智力。</li> <li>2. 學生人格。</li> <li>3. 父母對子女學業的態度。</li> <li>4. 社會或經濟地位。</li> </ol>
郭生玉 (1981)	<ol style="list-style-type: none"> <li>1. 心理因素：個人動機、人格適應、態度及學習習慣。</li> <li>2. 生理因素：視覺、聽覺及一般健康狀態。</li> <li>3. 社會因素：家庭背景、教育態度、家長職業及社區文化價值。</li> <li>4. 教育因素：教學方法、課程與教材等。</li> </ol>
王文科 (1991)	<ol style="list-style-type: none"> <li>1. 內在因素：指個人的智力、動機、人格特質</li> <li>2. 外在因素：指教學方法、學習團體、教師</li> </ol>
吳天泰 (1996)	個人因素：心理特質與環境
王財印 (2000)	<ol style="list-style-type: none"> <li>1. 身心方面因素，包括學生健康狀況、智力、性向、動機、適應等。</li> <li>2. 家庭方面因素，包括家庭環境、結構、管教方式。</li> <li>3. 學校方面因素，包括學校環境、教學內涵、教師特質與期望等。</li> <li>4. 社會方面因素，包括社區結構、文化、經濟等。</li> </ol>
陳怡華 (2001)	<ol style="list-style-type: none"> <li>1. 個人因素：個體的生理與心理等。</li> <li>2. 環境因素：家庭、學校教育與社會等。</li> </ol>
李美慧 (2004)	<ol style="list-style-type: none"> <li>1. 個人因素，包括自我概念、性別、動機、情緒、態度、智力等。</li> <li>2. 環境因素，包含家庭社經地位、父母親管教態度、重要他人的影響等。</li> </ol>
郭春悅 (2006)	<ol style="list-style-type: none"> <li>1. 個人因素，性別、智力、自我概念與學習動機。</li> <li>2. 環境因素，家庭與重要他人的影響。</li> </ol>

由上述研究整理中發現，影響學業成就的因素不外牽涉到個人、家庭、學校和社會等四個層面。本篇論文將從學生個人的學習背景因素的角度來探討其與學期成績間的關係。

## 第二節 資料採礦技術

資料採礦是利用統計以及機械學習的演算法，從大量資料中找尋隱藏具有商業價值的知識與規律，以作為自動化商業策略之應用。資料採礦是啟發性而非是演繹的，它揚棄了許多統計上的假設，啟發性的透過演算法主動搜尋有意義的規則，讓資料會說話(尹相志，民 98)。

在資料採礦的技術中依據採礦規則的型態可分為兩類：監督式學習(Supervised learning)與非監督式學習(Unsupervised learning)。監督式學習是依據輸入變數來找出與預測變數之間的關聯，因此這種採礦技術具備預測目標變數的能力。而非監督式學習主要是希望能從大量資料中找出一些隱藏的現象—資料間的相互關係，不一定要有預測的目標變數，而其目的在找出變數之間的關聯。

本論文所用的研究方法—關聯規則即屬於非監督式學習的一種。分析者在資料採礦的過程中發現規則，再由信心水準(Confidence)、支援度(Support)、重要性(Importance)等指標來判斷關聯規則的綜合成效(尹相志，民 98)。

資料採礦技術近年來被廣泛應用於各個領域，國內外許多的文獻及商業界都存在許多資料採礦成功的案例。例如：NCR 為美國最大零售商 Wal-Mart 所分析的案例，他們發現星期四晚上大多數購買尿布的消費者也會一併購買啤酒。所以，Wal-Mart 便在賣場將啤酒和尿布擺在一起，結果尿布和啤酒的銷售量雙雙成長。後來，研究發現原來這些同時購買尿布與啤酒的人大多是年輕的夫婦，他們趁星期四購買周末狂歡要用的啤酒時，也順便購買家中嬰兒的尿布。而這案例正突顯出利用資料採礦可以發掘出許多傳統方式與經驗無法找出的規律。

儘管資料採礦廣泛於各領域應用頗具成效，但資料採礦在教育界的研究卻相對稀少。其實對於資料成長量固定且保存完整的學生相關資料，若能有效發掘運用，定能使學校在行政及教學工作上有所貢獻(詹德麟，民 97)。

### 第三節 關聯規則

Agrawal 等學者於 1993 年提出資料採礦之關聯規則(Association Rules)方法。關聯規則在資料採礦的演算法中是一種可以產生大量規則的方法，使用者可以在大量的規則項目集合中找出感興趣的規則與關係。而關聯規則在商業上，主要是在找「哪些事情總是會同時發生」的規則，是一種零售業者常用來找出販售地商品之間的關聯性，以作為商品促銷及推薦的資料採礦技術，而這種技術有時也稱作購物籃分析(Basket Analysis) (尹相志，民 98)。

#### 壹、 關聯規則原理

關聯規則所運用的原理為條件機率，例如：購買 A 商品時，有多少機率會同時購買 B 商品。而關聯規則以信心水準 (Confidence)及支援度(Support)這兩個指標來評斷所找到的規則是否可用及是否有意義。

信心水準(在 SQL Server 2008 稱為機率 Probability)，是用來評斷所找到的關聯規則之準確度有多少，即在 A 的條件下發生 B 的機率有多少，如「公式 1」(尹相志，民 98)。

公式 1：信心水準

$$\text{Confidence (A} \Rightarrow \text{B)} = P(\text{A} \Rightarrow \text{B})$$

$$= P(\text{B}|\text{A})$$

$$= P(\text{A} \cap \text{B}) / P(\text{A})$$

支援度，是用來觀察所找到的關聯規則之案例支持情況，以比例的概念來呈現，即 A 與 B 同時發生的交易次數佔總交易次數的百分比來表示，如「公式 2」(尹相志，民 98)。亦可以數量的概念來呈現，即 A 與 B 同時發生的交易次數，稱為支援數，如「公式 3」(J. MacLennan et al., 2009)。

公式 2：支援度

$$\text{Support}(\text{A} \Rightarrow \text{B}) = P(\text{A} \cap \text{B})$$

公式 3：支援數

$$\text{Support}(\{\text{A}, \text{B}\}) = \text{Number of Transactions}(\text{A}, \text{B})$$

重要性(Importance)用來衡量所找到的關聯規則到底有無意義，即在有 A 的條件下發生 B 的機率，除以在無 A 的條件下發生 B 的機率，所得的商數再以 10 為底之對數來算出重要性指標，如「公式 4」。當重要性指標之值  $> 0$  代表正相關；若其值  $< 0$  代表負相關；而當值  $= 0$  代表無相關(J. MacLennan et al., 2009)。

公式 4：重要性

$$\text{Importance (A} \Rightarrow \text{B)} = \log ( P(\text{B}|\text{A}) / P(\text{B}|\text{not A}) )$$

「表 2-2」為早餐店的奶茶與蛋餅銷售關係表，以下將利用「表 2-2」來說明信心水準、支援及重要性指標的計算。

表 2-2 早餐店的豆漿與三明治銷售關係表

	買三明治	不買三明治	合計
買豆漿	32	12	44
不買豆漿	18	6	24
合計	50	18	68

$$\begin{aligned} \text{Support}(\text{買豆漿} \Rightarrow \text{買三明治}) &= 32 / 68 \\ &= 0.471 \end{aligned}$$

$$\text{Support}(\{\text{買豆漿}, \text{買三明治}\}) = 32$$

$$\begin{aligned} \text{Confidence}(\text{買豆漿} \Rightarrow \text{買三明治}) &= P(\text{買三明治} \mid \text{買豆漿}) \\ &= 32 / 44 \\ &= 0.727 \end{aligned}$$

$$\begin{aligned} \text{Importance}(\text{買豆漿} \Rightarrow \text{買三明治}) \\ &= \log(P(\text{買三明治} \mid \text{買豆漿}) / P(\text{買三明治} \mid \text{不買豆漿})) \\ &= \log((32 / 44) / (18 / 24)) \\ &= -0.013 \end{aligned}$$

由「表 2-2」早餐店豆漿與三明治銷售關係表計算（買豆漿 =>買三明治）的支援度、支援數、信心水準（Confidence）及重要性(Importance)指標。得到同時買豆漿與三明治的支援度是 0.471、支援數是 32；而在買豆漿的條件下同時買三明治的信心水準達 0.727；但由重要性指標為負值得知：在買豆漿的條件下會抑止同時買三明治的條件發生。

## 貳、 Apriori 演算法

關聯規則計算雖然容易，卻容易因為商品種類過多而造成規則大量膨脹。於是 Agrawal 與 Srikant 在 1994 年提出 Apriori 演算法，讓使用者在使用關連規則分析時可以快速地在大量的規則中搜尋出有效的規則。

Apriori 演算法是根據交易資料明細作為基礎，遞迴計算候選的物件組以及規則組合，茲利用「表 2-3」的交易資料為基礎說明 Apriori 演算法的運算流程。

表 2-3 交易資料表

發票編號	產品
001	A,B,E
002	A,C,D
003	B,C
004	A,B,C,E
005	A,C

步驟一：掃描「交易資料表」的全部交易紀錄，並計算各單一物件組的支援數後，產生「圖 2-1」中的「單 1 個物件的物件組」。

步驟二：以所設定的最小支援數(本例的最小支援數設為 3)對「單 1 個物件的物件組」中的物件組進行過濾，將支援數小於最小支援數的物件組予以刪除，得到「圖 2-1」中的「單 1 個物件的候選大型物件組」。



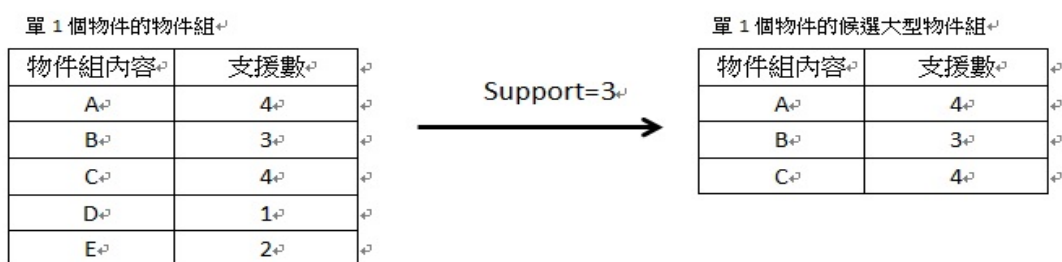


圖 2-1 資料表掃描與過濾結果(一)

步驟三：以「單 1 個物件的候選大型物件組」產生「圖 2-2」中的「2 個物件的物件組」(需去掉已刪除的物件組 D、E)，並統計其支援數。

步驟四：以所設定的最小支援數對「2 個物件的物件組」中的物件組進行過濾，將支援數小於最小支援數的物件組予以刪除，得到「圖 2-2」中的「2 個物件的候選大型物件組」。

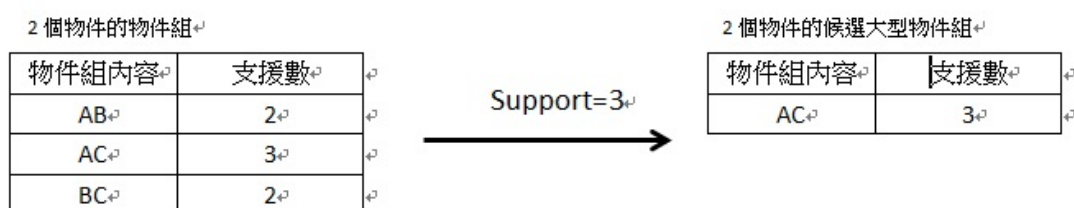


圖 2-2 資料表掃描與過濾結果(二)

步驟五：因為無法再以「2 個物件的候選大型物件組」產生「3 個物件的物件組」(需去掉已刪除的物件組 AB、BC)，至此完成 Apriori 演算法流程。

## 第三章 研究方法

本論文之研究方法採用資料採礦之關聯法則，分析國小學生個人學習背景因素對學期成績的預測能力。在第一節中，將學生原始成績與學習背景因素變項值調查結果做篩選處理並剔除不必要的部份，留下供資料採礦用的資料。在第二節中，將學生學期成績表與學習背景因素變項值調查結果分別轉換，成為資料採礦用的資料表。在第三節中，簡介資料採礦工具 SQL Server 2012。在第四節中，說明關聯規則篩選時之相關參數設定。

### 第一節 資料來源與資料篩選

本論文的資料是以雲林縣某國民小學之四、五、六年級普通班學生，其 102 學年度第一學期學期成績之平均與學生的個人學習背景因素變項值之問卷調查結果為研究樣本。排除未填問卷等遺漏資料後，得到學生學期成績資料與影響學生成績的個人學習背景因素資料 113 筆。

#### 壹、學生學期成績資料

學生學期成績資料為 102 學年度第一學期國語、英語、數學、自然、社會等五科的學期成績加總平均(小數點後四捨五入取整數)。再依據教育部『國民小學及國民中學學生成績評量準則』將學生學期成績的加總平均依五等第做成績的轉換。成績轉換等第對照表如「表 3-1」。學生學期成績如「表 3-2」。

表 3-1 成績轉換等第對照表

五等第	成績(分數)
優	90 分以上
甲	80 分以上未滿 90 分
乙	70 分以上未滿 80 分
丙	60 分以上未滿 70 分
丁	未滿 60 分

資料來源：教育部『國民小學及國民中學學生成績評量準則』

表 3-2 學生學期成績表

學生序號	性別	本國語文	英語	數學	自然與生 活科技	社會	學期平均	成績等第
4001	1	72	71	56	58	68	77	丙
4002	1	89	57	96	94	93	91	甲
4003	1	78	50	80	86	88	80	乙
4004	1	91	58	80	86	90	88	甲
4005	1	92	91	90	90	94	93	優
6043	2	78	41	57	74	80	81	丙
6044	2	70	70	51	72	79	82	丙
6045	2	98	97	97	92	95	95	優
6046	2	87	86	69	84	89	88	甲
6047	1	82	48	57	84	56	80	丙

## 貳、學生的個人學習背景因素

學生的個人學習背景因素變項值是透過線上問卷的方式

對四、五、六年級的普通班學生進行調查。問卷的題目、選

項改編自林建東(民 99)「影響學生成績的個人學習背景因素之問卷」詳如「表 3-3」。學生的個人學習背景因素變項值調查結果如「表 3-4」。

表 3-3 學生的個人學習背景因素之問卷與選項

題號	題目	選項	背景因素意義
1	請問你每天放學後，平均會花多少時間做功課、讀書（不包含課外補習的時間）？	(1)未達 1 個小時 (2)1 個小時以上～未達 2 個小時 (3)2 個小時以上～未達 3 個小時 (4)3 個小時以上	時間管理 (學習時間)
2	請問你每天放學後，平均會花多少時間從事休閒活動(例如：運動、看電視、聽音樂、讀課外書籍、使用電腦…)？	(1)未達 30 分鐘 (2)30 分鐘以上～未達 1 個小時 (3)1 個小時以上～未達 2 個小時 (4)2 個小時以上～未達 3 個小時 (5)3 個小時以上	時間管理 (休閒時間)
3	承上題，你每天的休閒活動中，平均有多少時間是在使用電腦、上網、玩電腦遊戲？	(1)未達 30 分鐘 (2)30 分鐘以上～未達 1 個小時 (3)1 個小時以上～未達 2 個小時 (4)2 個小時以上～未達 3 個小時 (5)3 個小時以上	時間管理 (使用電腦時間)
4	請問你每天在學校，平均利用多少下課時間做功課、讀書、準備考試，或與同學、老師討論功課？	(1)未達 20 分鐘 (2)20 分鐘以上～未達 30 分鐘 (3)30 分鐘以上～未達 40 分鐘 (4)40 分鐘以上	時間管理 (下課時間的利用)
5	請問你平均一星期有幾天會主動做功課、讀書？	(1)未達 1 天 (2)1 天 (3)2 天 (4)3 天 (5)3 天以上	學習動機 (主動學習)

(後續)

表 3-3 學生的個人學習背景因素之問卷與選項(續)

6	請問你這學期曾經安排過學習計畫，並依照計畫做功課、讀書嗎？	(1)不曾安排學習計畫 (2)有安排學習計畫，平均一週有未滿 1 天會依計畫做功課、讀書 (3)有安排學習計畫，平均一週有 1 天以上～未達 3 天會依計畫做功課、讀書 (4)有安排學習計畫，平均一週有 3 天以上～未達 5 天會依計畫做功課、讀書 (5)有安排學習計畫，平均一週有 5 天以上會依計畫做功課、讀書	學習計畫與執行力
7	請問你平均一星期有幾天做功課、讀書與晚上睡覺、早上起床的時間是固定的？	(1)未達 1 天 (2)1 天 (3)2 天 (4)3 天 (5)3 天以上	作息規律性
8	請問你平日讀書、做功課遇到問題，自己無法解決時，你會找家人、朋友、同學或老師來幫忙解決問題嗎？	(1)平均 10 次中有 8 次以上會找上述的人幫忙解決問題 (2)平均 10 次中有 5 次以上～未滿 8 次會找上述的人幫忙解決問題 (3)平均 10 次中有 2 次以上～未滿 5 次會找上述的人幫忙解決問題 (4)平均 10 次中有未滿 2 次會找上述的人幫忙解決問題	虛心求助頻率
9	請問你目前參加校外補習或是請家教指導的科目有幾科？	(1)無 (2)1 科 (3)2 科 (4)3 科以上	校外補習科目數
10	請問在本學期中，你很認真學習的科目(國語、英語、數學、自然、社會)有幾個？	(1)無 (2)1 科 (3)2 科 (4)3 科以上	學習態度

表 3-4 學生個人學習背景因素變項值調查結果

學生序號	學習時間	休閒時間	使用電腦 時間	下課時間 的利用	主動學習	學習計畫 與執行力	作息 規律性	虛心求助 頻率	校外補習 科目數	學習態度
4001	4	1	2	1	2	1	2	1	1	4
4002	4	5	2	2	5	5	5	4	2	4
4003	4	2	1	4	1	5	5	3	1	4
4004	2	5	4	1	1	2	2	4	4	1
4005	3	3	2	2	5	1	5	1	1	4
4006	3	2	1	1	5	3	5	1	1	4
4007	1	2	2	1	4	1	4	1	1	4
4008	2	1	1	1	5	5	5	4	1	4
4009	1	3	4	1	2	1	3	1	1	3
4010	2	2	2	1	3	1	5	1	2	3

## 第二節 資料的預處理

將「表 3-2」學生學期成績表與「表 3-4」學生個人學習背景因素變項值調查結果分別轉換成「表 3-5」學生學期成績主表及「表 3-6」學生個人學習背景因素資料表(以二筆資料為例)，做為資料採礦用的資料表：

表 3-5 學生學期成績主表

學生序號	成績等第
4001	丙
4002	甲

表 3-6 學生個人學習背景因素資料表

學生序號	變項	變項值
4001	學習時間	4
4001	休閒時間	1
4001	使用電腦時間	2
4001	下課時間的利用	1
4001	主動學習	2
4001	學習計畫與執行力	1
4001	作息規律性	2
4001	虛心求助頻率	1
4001	校外補習科目數	1
4001	學習態度	4
4002	學習時間	4
4002	休閒時間	5
4002	使用電腦時間	2
4002	下課時間的利用	2
4002	主動學習	5
4002	學習計畫與執行力	5
4002	作息規律性	5
4002	虛心求助頻率	4
4002	校外補習科目數	2
4002	學習態度	4

### 第三節 資料採礦工具 SQL Server 2012 簡介

Microsoft SQL Server 是由美國微軟公司所推出的關聯式資料庫解決方案，最新的版本是 SQL Server 2012，已經在 2012 年 3 月 6 日發行。SQL Server 2012 提供了一組多樣化的整合式服務，能對資料進行如查詢、搜尋、同步化、報告與分析等多種操作。資料可以儲存在資料中心中的最大型伺服器上，也能儲存在桌上型電腦與行動裝置，無論儲存位置為何，都能充分控制資料。

SQL Server 2012 能讓使用者利用 Microsoft.NET 與 Visual Studio 開發的客製化應用程式，以及經由 Microsoft BizTalk Server 內的服務導向架構(SOA) 與業務程序來存取資料，而且資訊工作者也能使用工具直接存取每日使用的資料，如 Microsoft Office 2007 System。SQL Server 2012 提供高信賴度、高生產力與商業智慧的資料平台，能符合客戶所有的資料需求。

Microsoft Visual Studio (簡稱 VS) 是美國微軟公司的開發工具套件系列產品。VS 是一個基本完整的開發工具集，具有 SQL Server 商務智慧特有的其他專案類型。提供開發包含 Analysis Services、Integration Services 和 Reporting Services 專案的商務方案。每一個專案類型提供建立商務智慧方案之必要物件的範本，並提供各種設計師、工具和精靈來使用物件。

本篇論文採用 Microsoft Visual Studio 2010 為資料採礦工具，其演算法計有：決策樹、貝氏機率分類、時序群集、時間序列、群集、線性迴歸、羅吉斯迴歸、類神經網路、關聯規則等九種演算法。



#### 第四節 演算法參數設定

本論文研究方法採用 SQL Server 2012 資料庫搭配 Microsoft Visual Studio 2010 資料採礦之關聯法則，其系統的演算法參數設定如表 3-7 系統的演算法參數表。

表 3-7 系統的演算法參數表

參數項目	參數值
MAXIMUM_ITEMSET_SIZE	0
MINIMUM_PROBABILITY	0.3
MINIMUM_SUPPORT	6
MINIMUM_IMPORTANCE	0.1

以下將依序說明操作過程：

- 一、本研究資料採礦用的資料表共有 10 個變項，MAXIMUM\_ITEMSET\_SIZE(項目集內所允許的最大項目數目) 設為 0，代表本研究項目集沒有大小限制。
- 二、先分別計算出學生各成績等第的人數，如「表 3-8」。

表 3-8 學生各成績等第的人數分配表

等第	人數
優	29
甲	36
乙	24
丙	14
丁	10

三、由表 3-8 學生各成績等第的人數分配表得知，學期成績等第「丁」是分配表個數最少的項目，故以小於等第「丁」的次數 6 來作為關聯規則的 MINIMUM\_SUPPORT(最小支援數)。

四、將 MINIMUM\_PROBABILITY(最小信心水準)設為 0.3，MINIMUM\_IMPORTANCE(最小重要性)設為 0.1，從採礦模型檢視器的規則中篩選出 15 條以內的規則。如果篩選出的規則數量不在此一範圍，調整最小信心水準及最小重要性，重新進行採礦結構程序之處理，直至找到符合的規則。

## 第四章 結果與討論

本章共分三節，第一節為資料特性分析，對取得的資料先做一敘述性統計，整理成各成績等第人（次）數分佈表，綜合說明觀察的現象；第二節說明關聯法則模型的建置過程；第三節為資料採礦-關聯規則的彙整與分析。

### 第一節 資料特性分析

壹、性別資料分析：

113 筆資料中，男生佔 54 人，女生佔 59 人。從「圖 4-1 學生性別比例分佈圖」中可看出比例約 48：52，人數差距約 4 %。

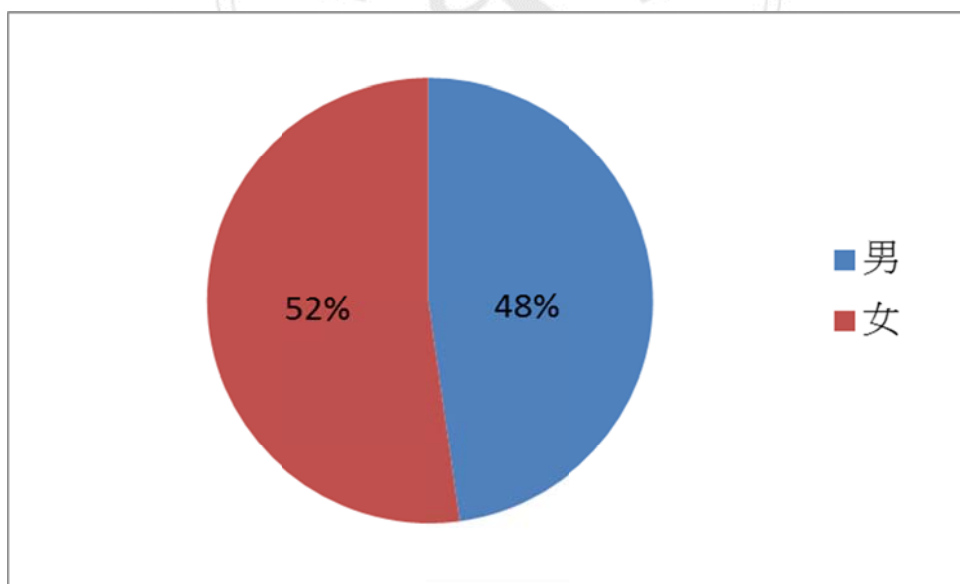


圖 4-1 學生性別比例分佈圖

貳、以學期成績等第與性別作增益值分析：

增益值，就是將子體發生機率除以母體的原始發生機率所得到的比值。增益值 $>1$ 表示發生機率高於母體。因此，該條件會被視作為傾向發生。

從「表 4-1 學期成績等第人數分佈表」觀察、對照後，結果發現：等第-優部份，女生人數增益值為 1.25，其值大於 1 代表該條件會傾向發生，代表得「優」的女生相對較多；等第-甲部份，男生人數增益值為 1.1，代表得「甲」的男生相對較多；等第-乙部份，男生人數增益值為 1.22，代表得「乙」的男生相對較多。等第-丙部份，男生人數增益值為 1.19，代表得「丙」的男生相對較多。等第-丁部份，女生人數增益值為 1.34，代表得「丁」的女生相對較多。

表 4-1 學期成績等第人數分佈表

成績等第	優	甲	乙	丙	丁		
分組總人數	29	36	24	14	10	113	全體總人數
男生人數	10	19	14	8	3	54	男生總人數
女生人數	19	17	10	6	7	59	女生總人數
男生增益值	0.72	1.1	1.22	1.19	0.63		
女生增益值	1.25	0.9	0.79	0.82	1.34		

## 第二節 關聯法則模型建置過程

### 壹、 連接資料採礦資料庫

步驟一：開啟 SQL Server Management Studio 連結資料庫引擎後，新增一個名為 ScoreDMdb 的資料庫，如圖 4-2 新增資料庫 ScoreDMdb。

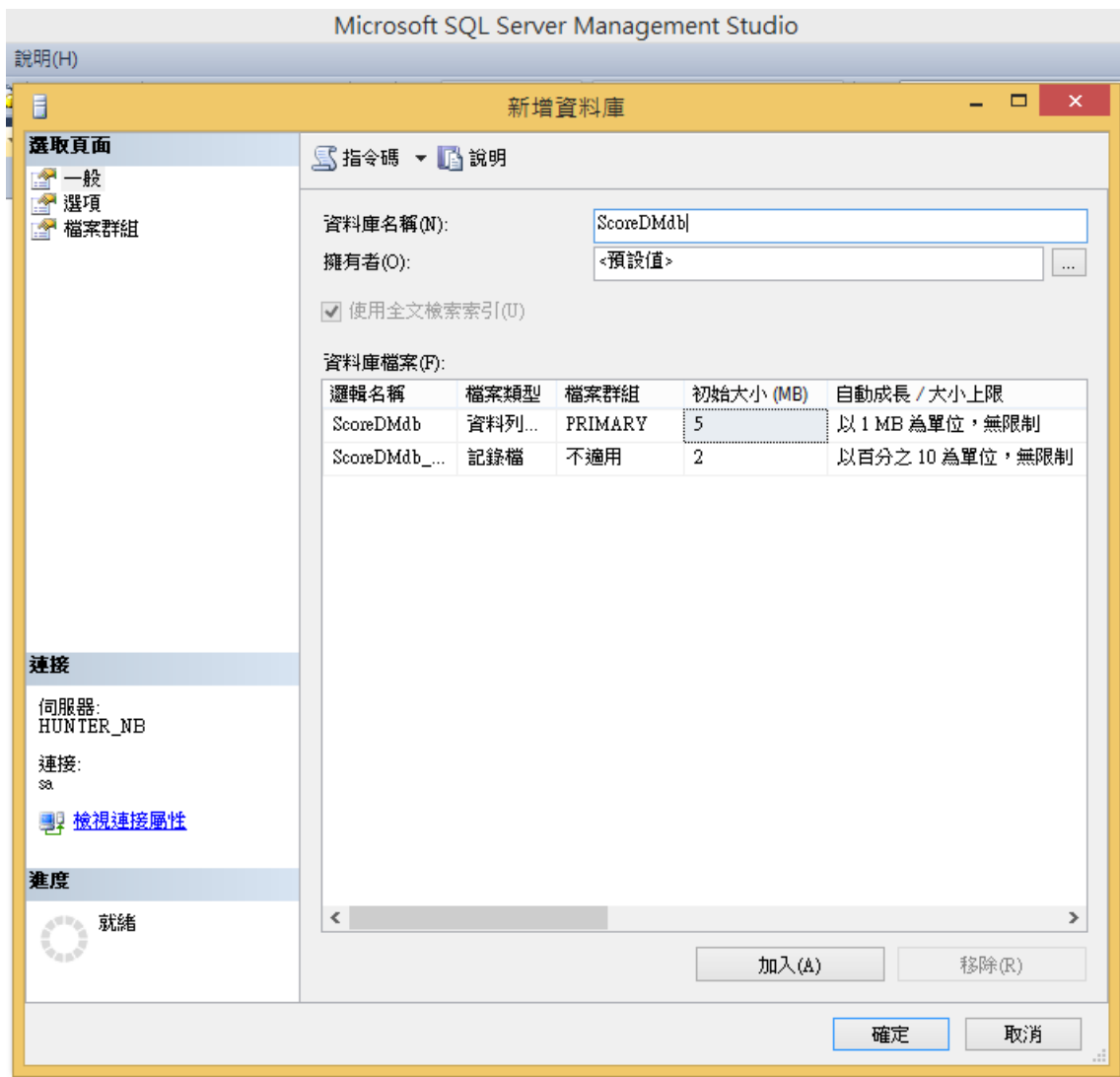


圖 4-2 新增資料庫 ScoreDMdb

步驟二：選取 ScoreDMdb 資料庫，按滑鼠右鍵選擇「工作」→「匯入資料」，開啟 SQL Server 匯入和匯出精靈，

資料來源選擇 Microsoft Excel 格式，選擇檔案路徑後分別匯入 Excel 資料表 Scoremaintb.xlsx 及 Scorefactortb.xlsx，如圖 4-3 匯入 Scorefactortb.xlsx 資料表。

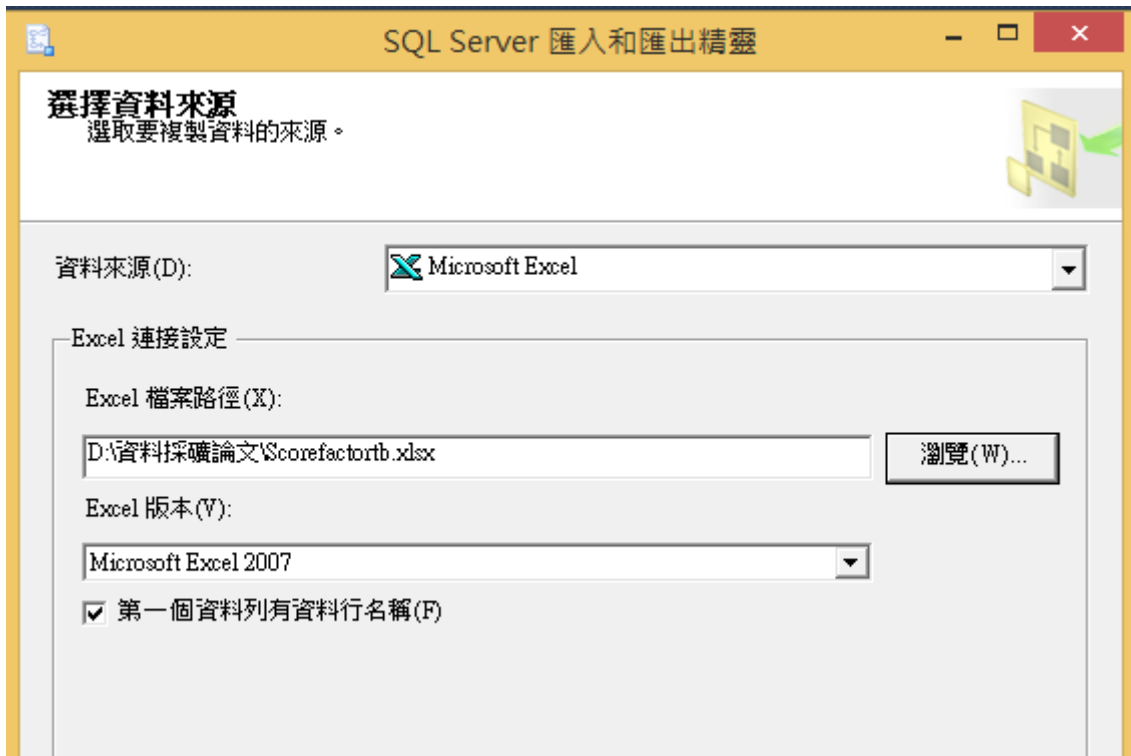


圖 4-3 匯入 Scorefactortb.xlsx 資料表

步驟三：選取來源資料表 Scoremain\$ 及 Scorefactor\$，檢視並設定好資料表內適當的資料類型，如圖 4-4 設定 scorefactor\$ 資料表適當的資料類型。成功連結後按「關閉」。

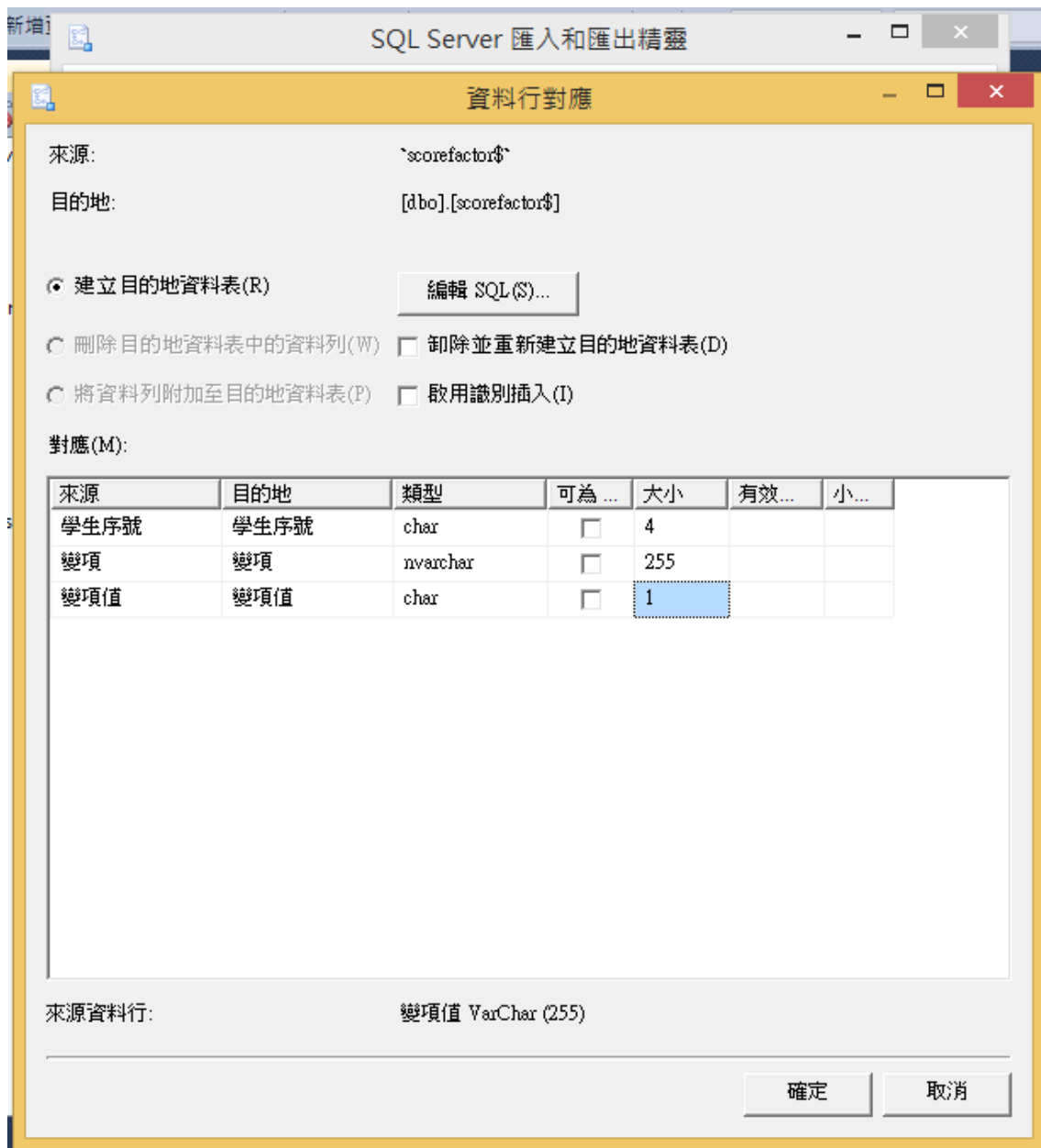


圖 4-4 設定 scorefactor\$資料表適當的資料類型

步驟四：分別將 scoremain\$資料表的「學生序號」及 scorefactor\$資料表的「學生序號」和「變項」設定為主鍵(旁有鑰匙符號)，如圖 4-5、圖 4-6，至此完成整個連結 ScoreDMdb 資料庫的流程。

HUNTER_NB.Score...dbo.scoremain\$*		
資料行名稱	資料類型	允許 Null
學生序號	char(4)	<input type="checkbox"/>
成績等第	nvarchar(1)	<input type="checkbox"/>
		<input type="checkbox"/>

圖 4-5 設定主鍵畫面一

HUNTER_NB.Score...dbo.scorefactor\$*			HUNTER_NB.Score...db
資料行名稱	資料類型	允許 Null	
學生序號	char(4)	<input type="checkbox"/>	
變項	nvarchar(255)	<input type="checkbox"/>	
變項值	char(1)	<input type="checkbox"/>	
		<input type="checkbox"/>	

圖 4-6 設定主鍵畫面二

## 貳、建置關聯規則的採礦結構

步驟一：開啟 SQL Server Data Tools，建立一個 Analysis Services 多維度和資料採礦專案命名為 ScoreDMProj，如圖 4-7 新增採礦專案 ScoreDMProj。



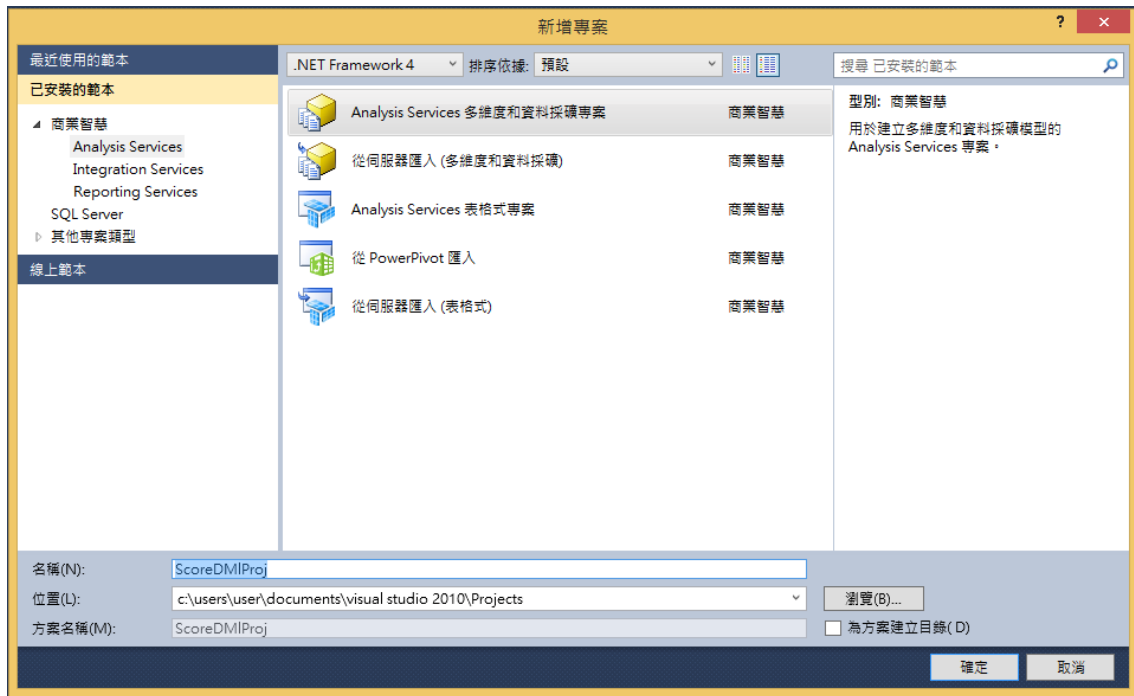


圖 4-7 新增採礦專案 ScoreDMProj

步驟二：開啟 ScoreDMProj 專案，點選「方案總管」內的「資料來源」，按滑鼠右鍵選擇「新增資料來源」如圖 4-8 ScoreDMProj 專案新增資料來源；接著連接之前建立的「ScoreDMdb」資料庫，如圖 4-9 ScoreDMProj 專案連接 ScoreDMdb 資料庫，完成資料來源的設定。

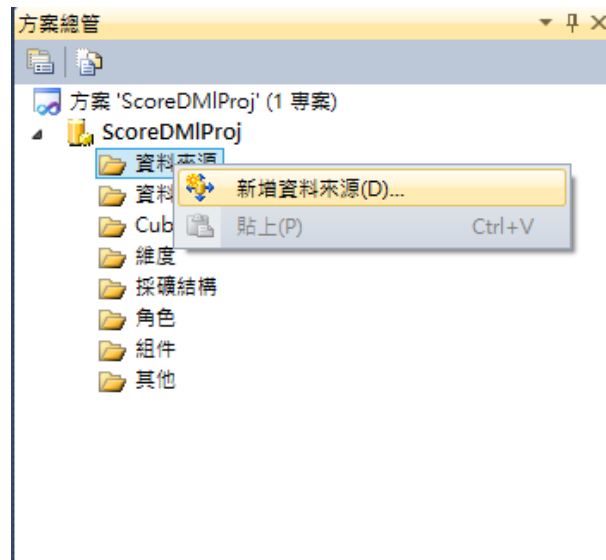


圖 4-8 ScoreDMProj 專案新增資料來源

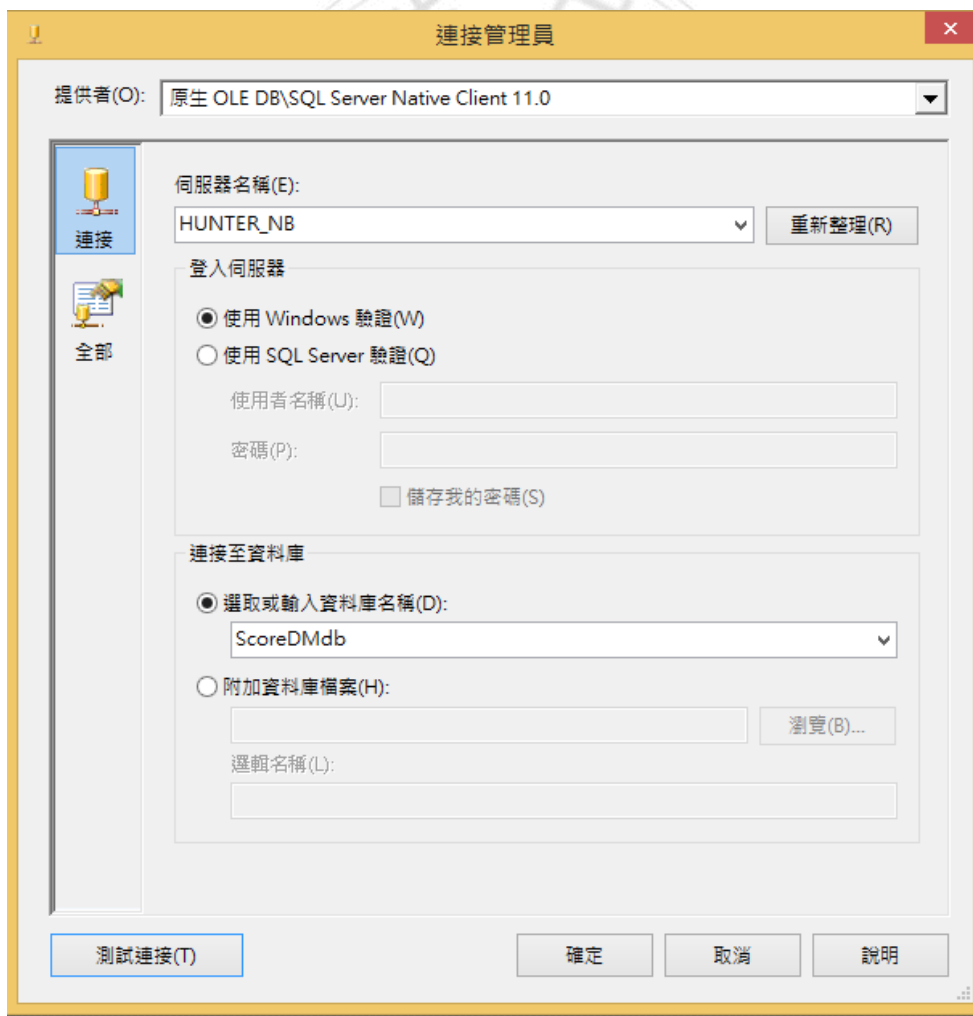


圖 4-9 ScoreDMProj 專案連接 ScoreDMdb 資料庫

步驟三：點選「方案總管」內的「資料來源檢視」按滑鼠右鍵選擇「新增資料來源檢視」如圖 4-10 ScoreDMProj 專案新增資料來源檢視；選取並連接名為 scoremain\$ 和 scorefactor\$ 資料表的關聯式資料來源，如圖 4-11 ScoreDMProj 專案連接 scoremain\$ 和 scorefactor\$ 資料表，完成資料來源檢視的設定。

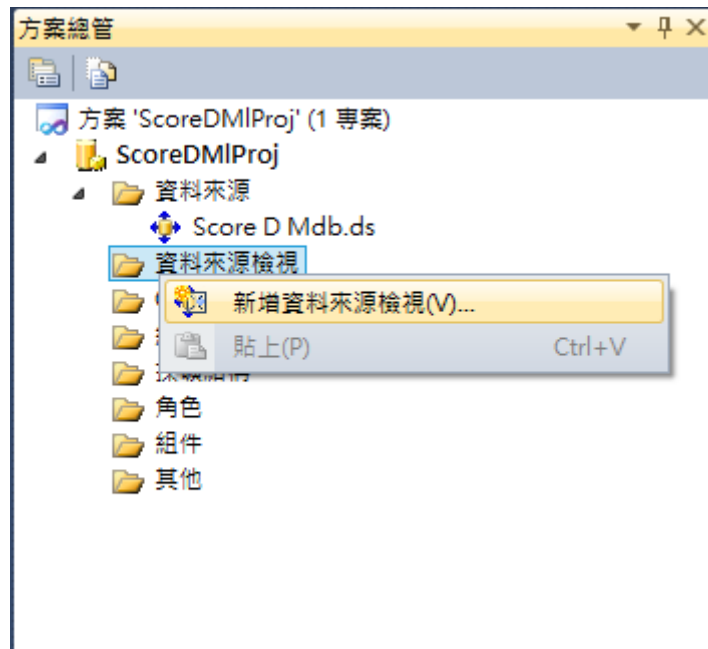


圖 4-10 ScoreDMProj 專案新增資料來源檢視

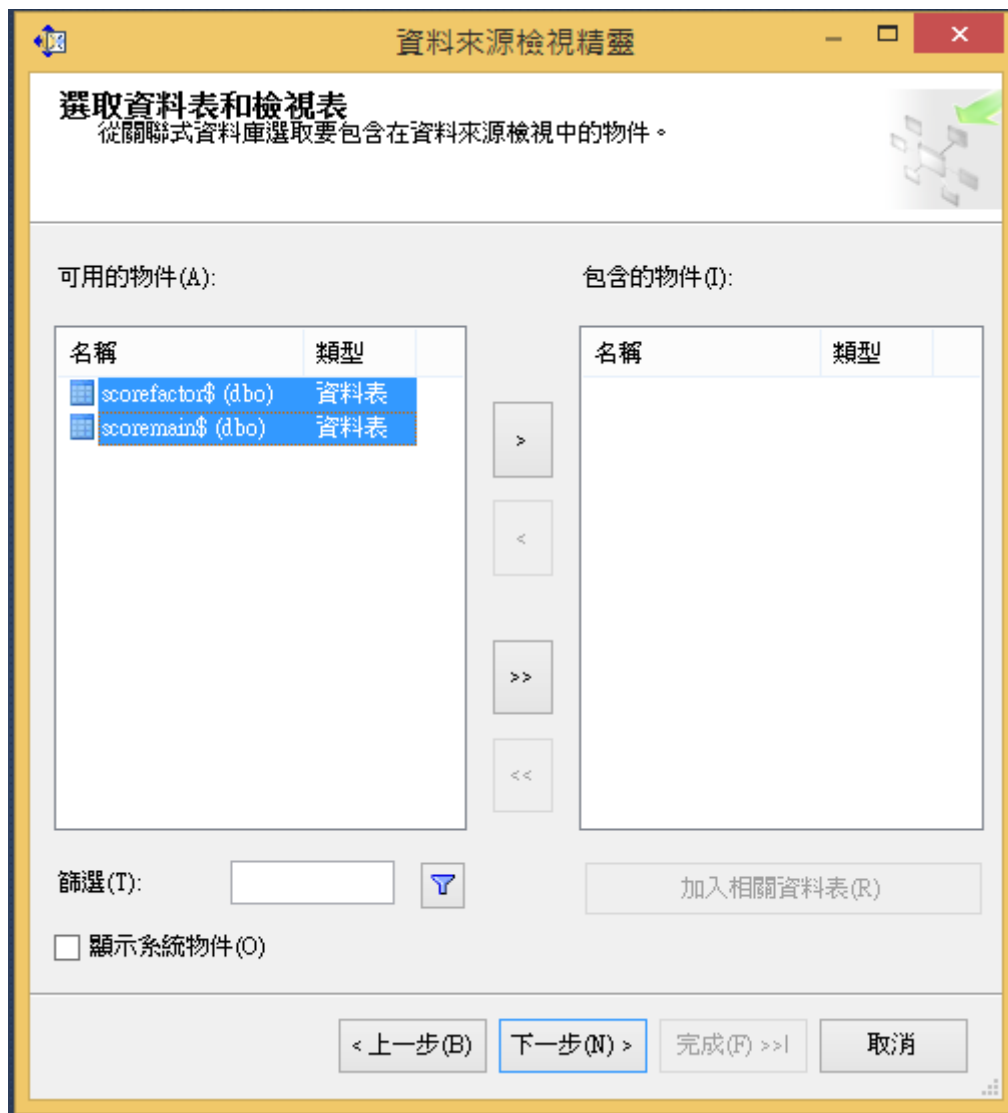


圖 4-11 ScoreDMProj 專案連接 scoremain\$和 scorefactor\$資料表  
 步驟四：在 scorefactor\$資料表新增具名計算「變項與變項  
 值組合」，如圖 4-12。

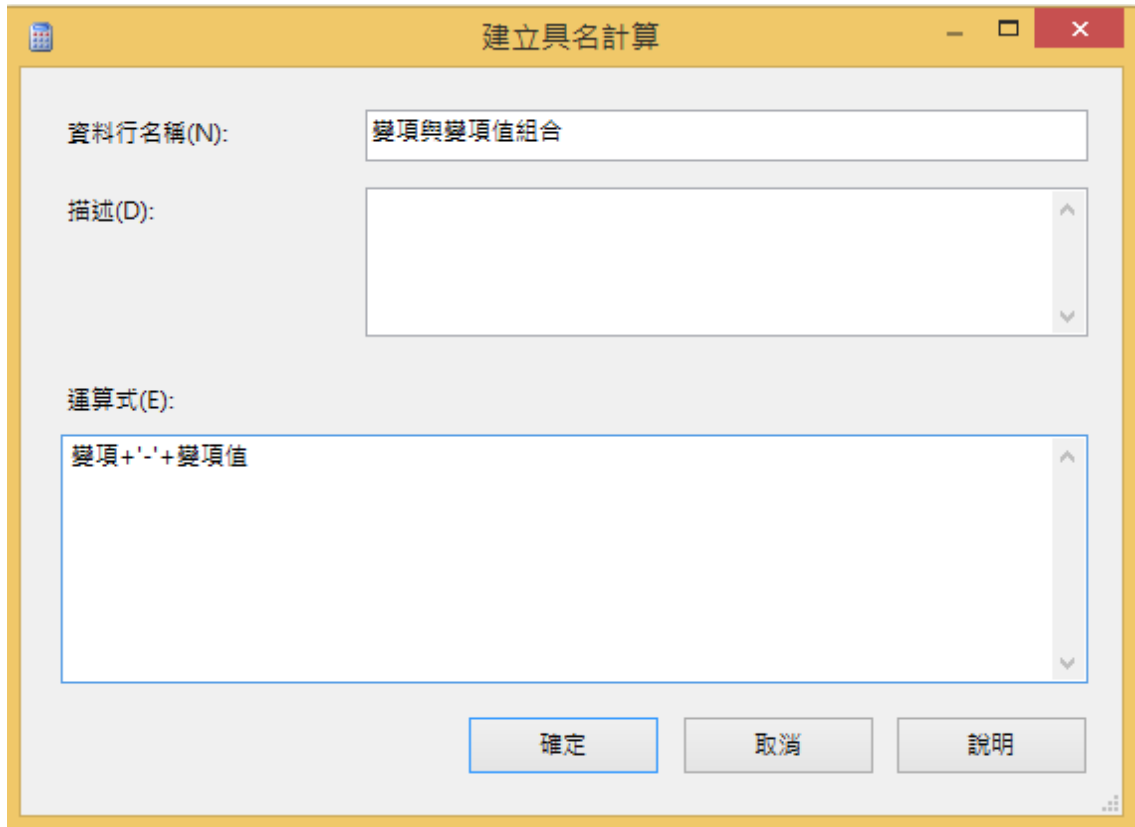


圖 4-12 scorefactor\$資料表新增具名計算

步驟五：點選「方案總管」內的「採礦結構」，按滑鼠右鍵選擇「新增採礦結構」（如圖 4-13），選取「Microsoft 關聯規則」資料採礦技術（如圖 4-14）。

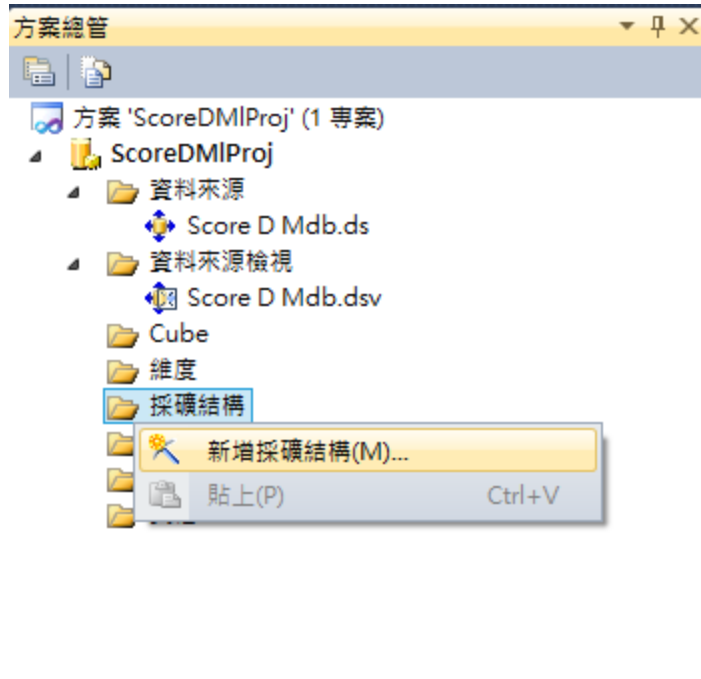


圖 4-13 新增採礦結構

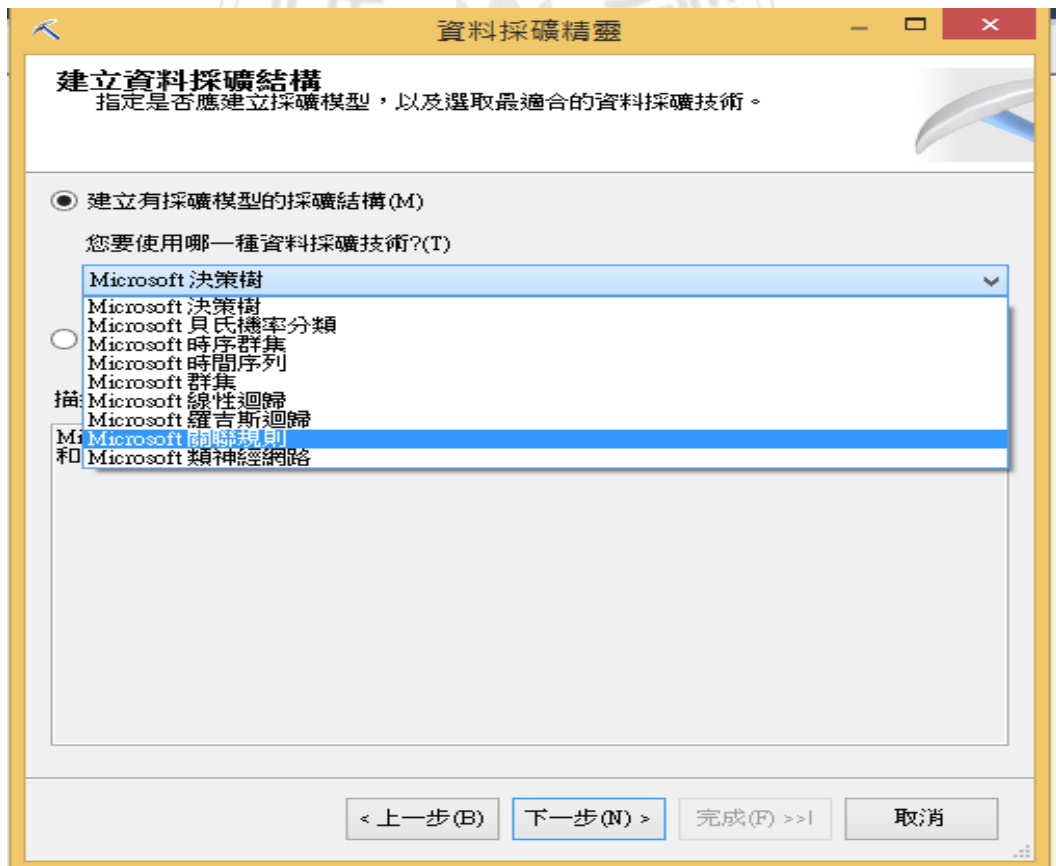


圖 4-14 選取 Microsoft 關聯規則技術

步驟六：選取資料來源檢視裡的 scoremain\$和 scorefactor\$資料表，在「scoremain\$資料表」裡勾選「案例」選項，在「scorefactor\$資料表」裡勾選「巢狀」選項(如圖 4-15)。

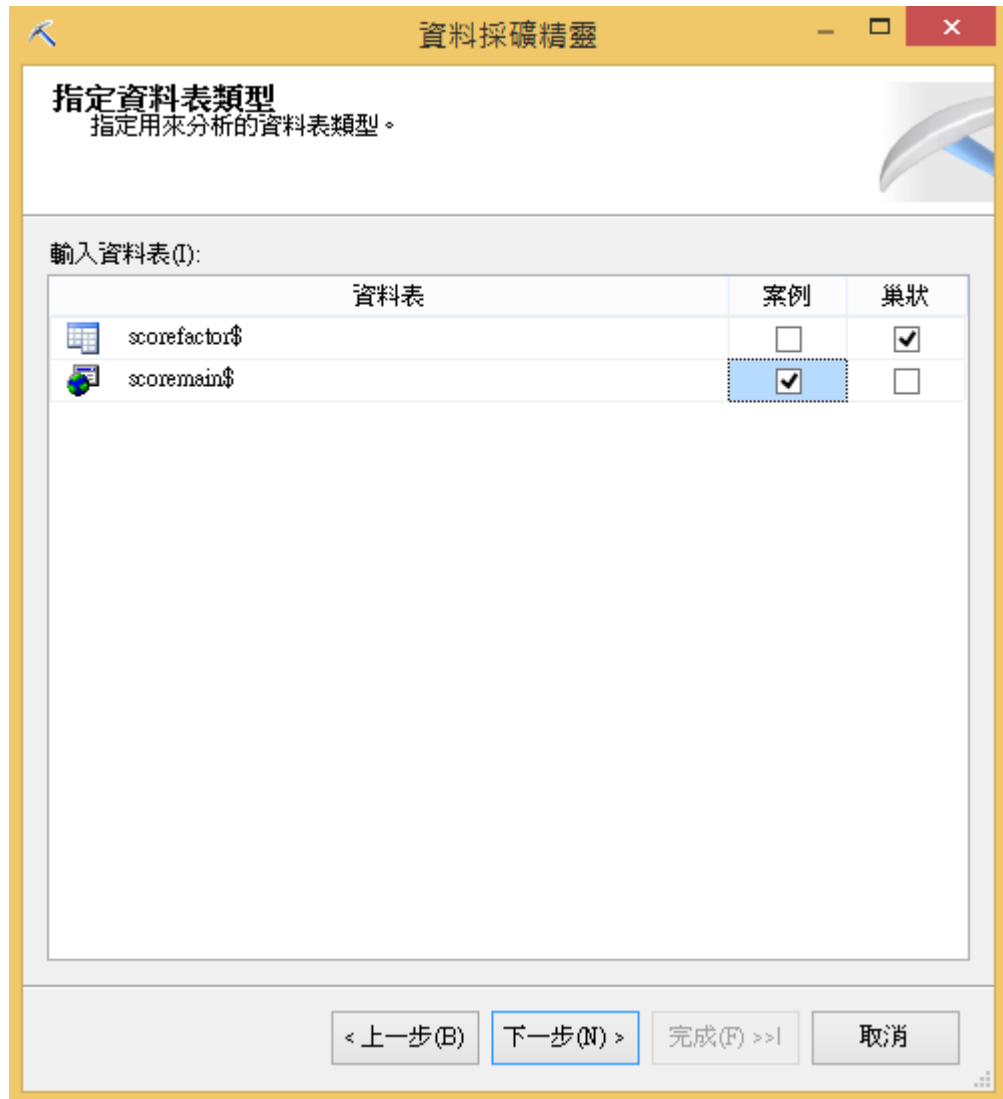


圖 4-15 指定 scoremain\$和 scorefactor\$資料表類型

步驟七：設定好訓練資料的索引鍵、輸入和可預測選項(如圖 4-16)，確定資料類型為「類別變數」(如圖 4-17)，

即完成新增採礦結構的動作。接著，就可以開始第三節規則的尋找過程。



圖 4-16 指定分析使用的資料行



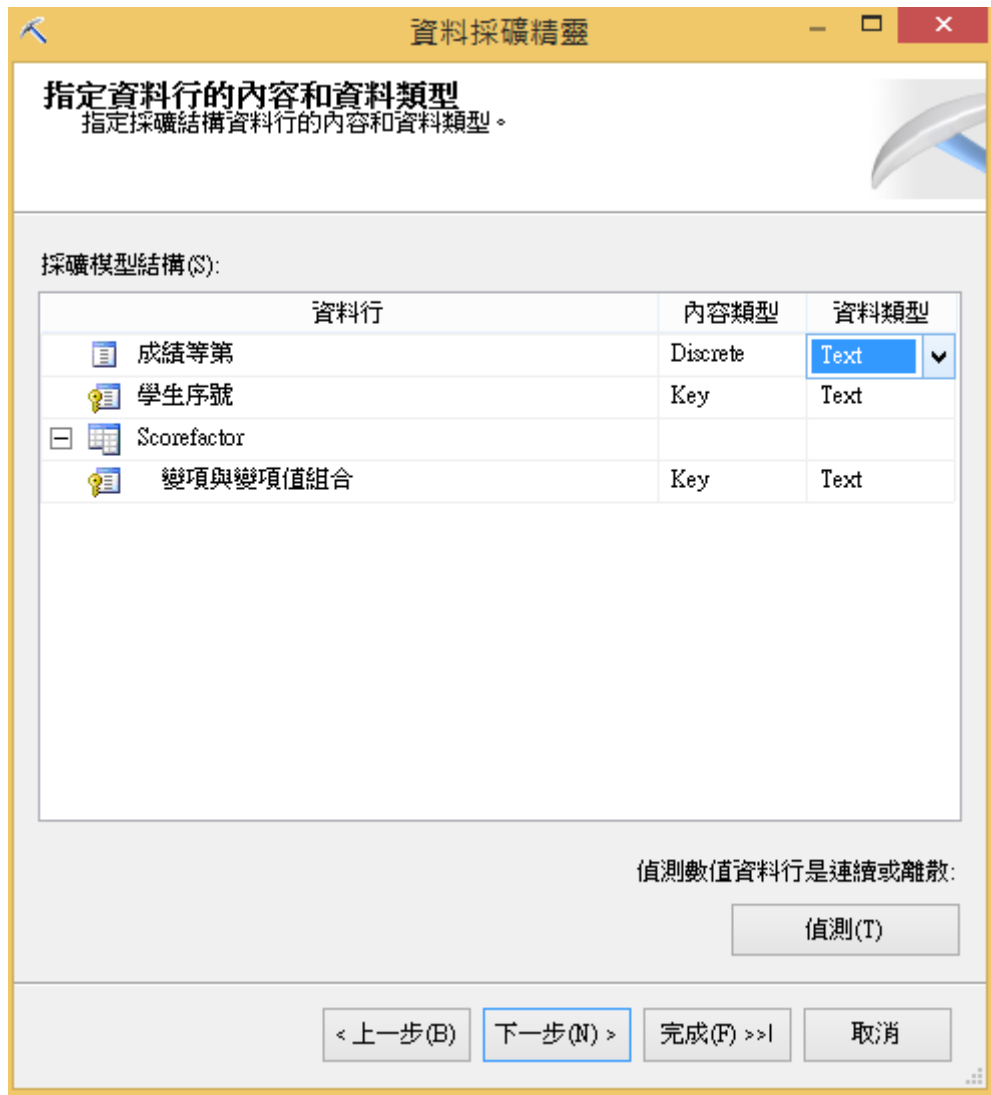


圖 4-17 指定分析資料行的資料類型為「類別變數」

### 參、設定系統演算法參數

依據第三章第四節設定演算法參數，如圖 4-18 關聯規則演算法參數設定圖。

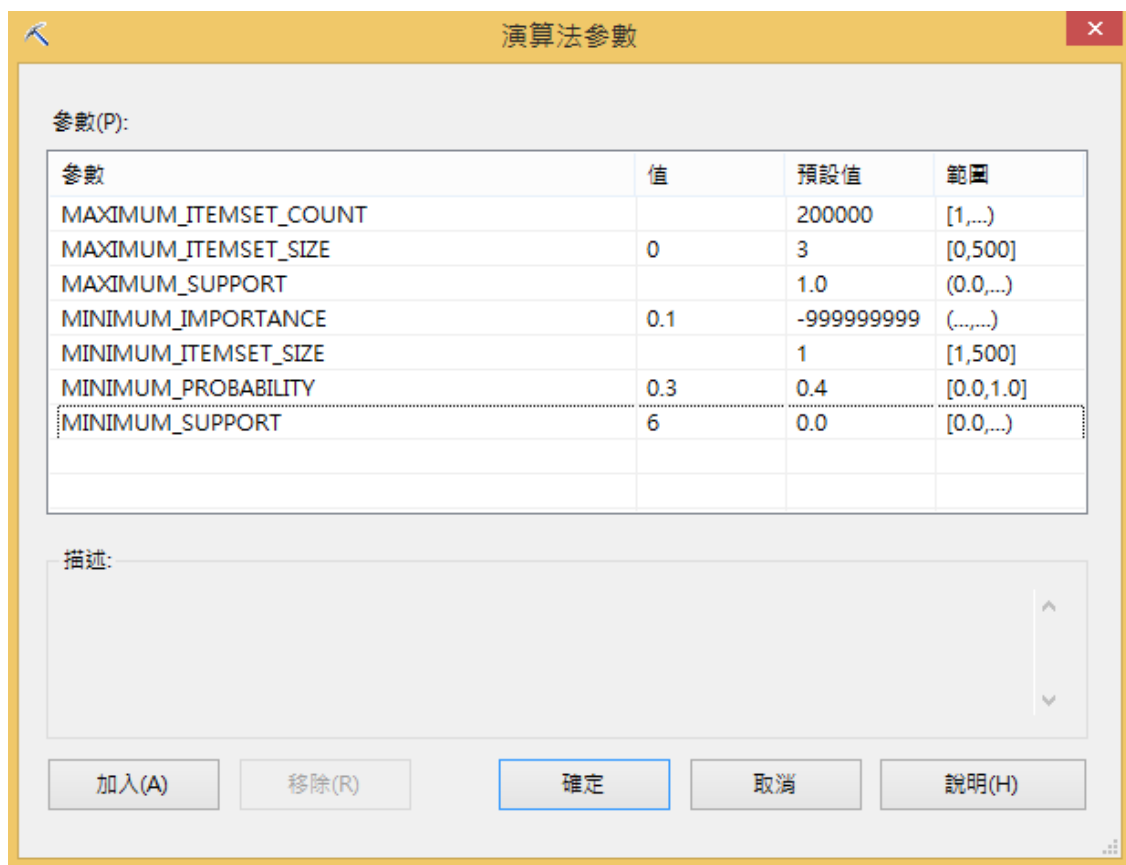


圖 4-18 設定關聯規則演算法參數

### 第三節 資料採礦-關聯規則的彙整與分析

本論文要分析的是「學生個人學習背景因素對國小學業成就的預測能力」，因此變項與變項值的組合成為輸入變數以及巢狀資料表索引鍵，而成績等第成為「可預測變數」（關聯規則的箭頭後方只會出現成績等第變數）。「可預測變數」有五種情況：「成績等第=優」、「成績等第=甲」、「成績等第=乙」、「成績等第=丙」及「成績等第=丁」。

#### 壹、關聯規則彙整

##### 一、「成績等第=優」規則彙整

「篩選規則」輸入「成績等第=優」，以圖 4-18 預設系統演算法參數執行採礦程序，符合規則有 67 條，超出 15

條以內的範圍，於是反覆調整各項參數，尋找符合的規則，較佳參數如表 4-2：

表 4-2 「成績等第=優」參數設定表

最小支援數	最小機率	最低重要性	最大項目集
6	0.61	0.4	不限

以表 4-2 的參數執行採礦程序得規則 15 條，符合研究所需，如附錄一的圖 1，整理規則如表 4-3，依重要性欄位由大至小排序以利規則分析，相依性網路如圖 4-19 所示。

表 4-3 「成績等第=優」規則彙整表

機率	重要性	影響因素組合
0.818	0.580	使用電腦時間-2, 主動學習-5, 學習態度-4
0.733	0.570	使用電腦時間-2, 主動學習-5
0.857	0.544	使用電腦時間-2, 主動學習-5, 學習態度-4, 作息規律性-5
0.857	0.544	虛心求助頻率-2, 主動學習-5, 學習態度-4
0.857	0.544	學習計畫與執行力-4, 使用電腦時間-2
0.778	0.525	使用電腦時間-2, 學習態度-4, 作息規律性-5
0.750	0.494	休閒時間-3, 主動學習-5, 學習態度-4
0.750	0.494	學習時間-3, 主動學習-5, 學習態度-4
0.700	0.483	使用電腦時間-2, 主動學習-5, 作息規律性-5
0.700	0.483	學習時間-3, 主動學習-5
0.700	0.483	虛心求助頻率-2, 主動學習-5
0.667	0.449	學習計畫與執行力-4
0.667	0.449	虛心求助頻率-2, 學習態度-4
0.667	0.449	學習計畫與執行力-1, 主動學習-5, 學習態度-4
0.667	0.449	校外補習科目數-4, 主動學習-5, 學習態度-4

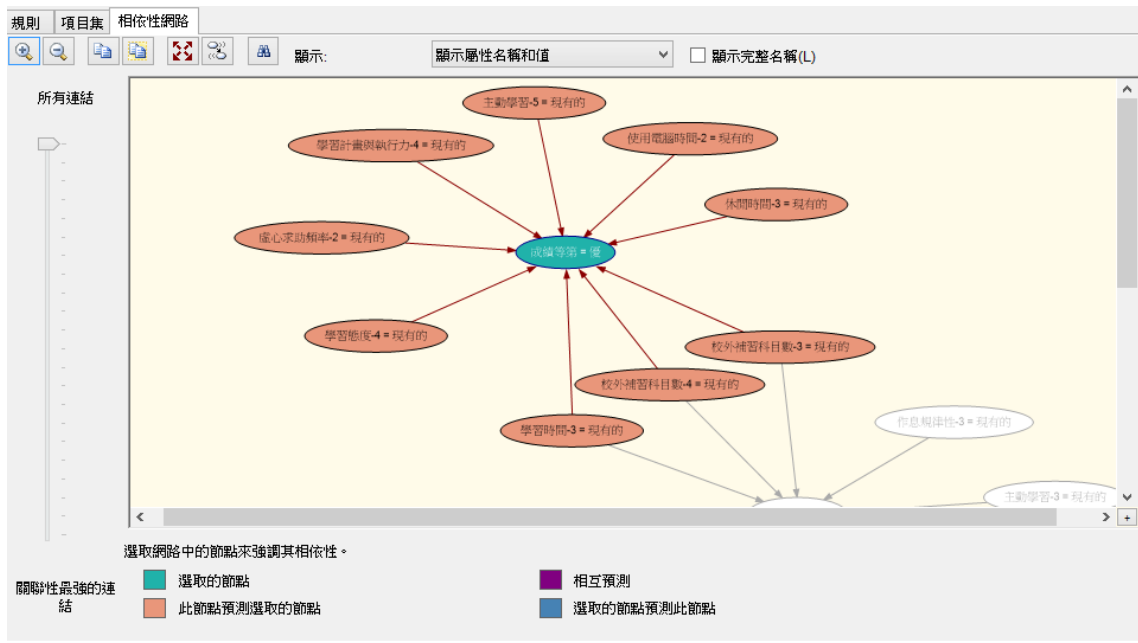


圖 4-19 「成績等第=優」相依性網路圖

## 二、「成績等第=甲」規則彙整

「篩選規則」輸入「成績等第=甲」，以圖 4-18 預設系統演算法參數執行採礦程序，符合規則有 42 條，超出 15 條以內的範圍，於是反覆調整各項參數，尋找符合的規則，較佳參數如表 4-4：

表 4-4 「成績等第=甲」參數設定表

最小支援數	最小機率	最低重要性	最大項目集
6	0.41	0.2	不限

以表 4-4 的參數執行採礦程序得規則 13 條，符合研究所需，如附錄一的圖 2，整理規則如表 4-5，依重要性欄位由大至小排序以利規則分析，相依性網路如圖 4-20 所示。

表 4-5 「成績等第=甲」規則彙整表

機率	重要性	影響因素組合
0.875	0.455	校外補習科目數-4, 學習時間-2
0.857	0.433	校外補習科目數-4, 下課時間的利用-2
0.750	0.383	學習計畫與執行力-3, 使用電腦時間-1, 虛心求助頻率-1
0.727	0.395	學習計畫與執行力-3, 使用電腦時間-1
0.727	0.395	學習時間-2, 虛心求助頻率-1, 作息規律性-5
0.667	0.338	主動學習-3, 虛心求助頻率-1
0.667	0.338	學習時間-2, 學習態度-4, 虛心求助頻率-1
0.600	0.296	學習時間-3, 休閒時間-3
0.600	0.296	使用電腦時間-1, 學習態度-4, 虛心求助頻率-1
0.583	0.293	下課時間的利用-2, 學習態度-4
0.500	0.252	學習時間-2, 虛心求助頻率-1
0.500	0.226	校外補習科目數-3, 作息規律性-5
0.500	0.220	下課時間的利用-2, 虛心求助頻率-1

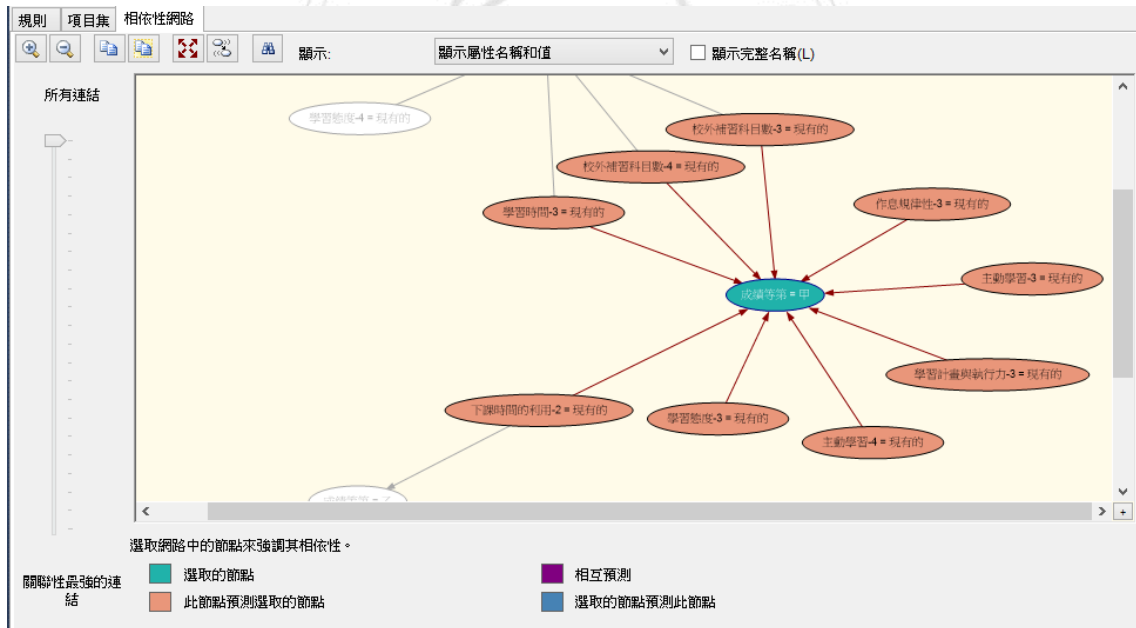


圖 4-20 「成績等第=甲」相依性網路圖

### 三、「成績等第=乙」規則彙整

「篩選規則」輸入「成績等第=乙」，以圖 4-18 預設系統演算法參數執行採礦程序，符合規則有 11 條，符合研究

所需，如附錄一的圖 3，整理規則如表 4-6，依重要性欄位由大至小排序以利規則分析，相依性網路如圖 4-21 所示。

表 4-6 「成績等第=乙」規則彙整表

機率	重要性	影響因素組合
0.636	0.551	下課時間的利用-2, 學習計畫與執行力-3
0.538	0.480	校外補習科目數-1, 主動學習-5, 作息規律性-5
0.500	0.464	校外補習科目數-1, 主動學習-5
0.462	0.399	學習態度-3, 學習計畫與執行力-1
0.409	0.384	校外補習科目數-1, 作息規律性-5
0.429	0.367	下課時間的利用-2, 學習態度-3
0.368	0.308	校外補習科目數-1, 學習態度-4
0.345	0.308	下課時間的利用-2
0.375	0.307	下課時間的利用-2, 學習時間-2
0.350	0.283	學習計畫與執行力-1, 虛心求助頻率-1, 下課時間的利用-1
0.320	0.247	使用電腦時間-3

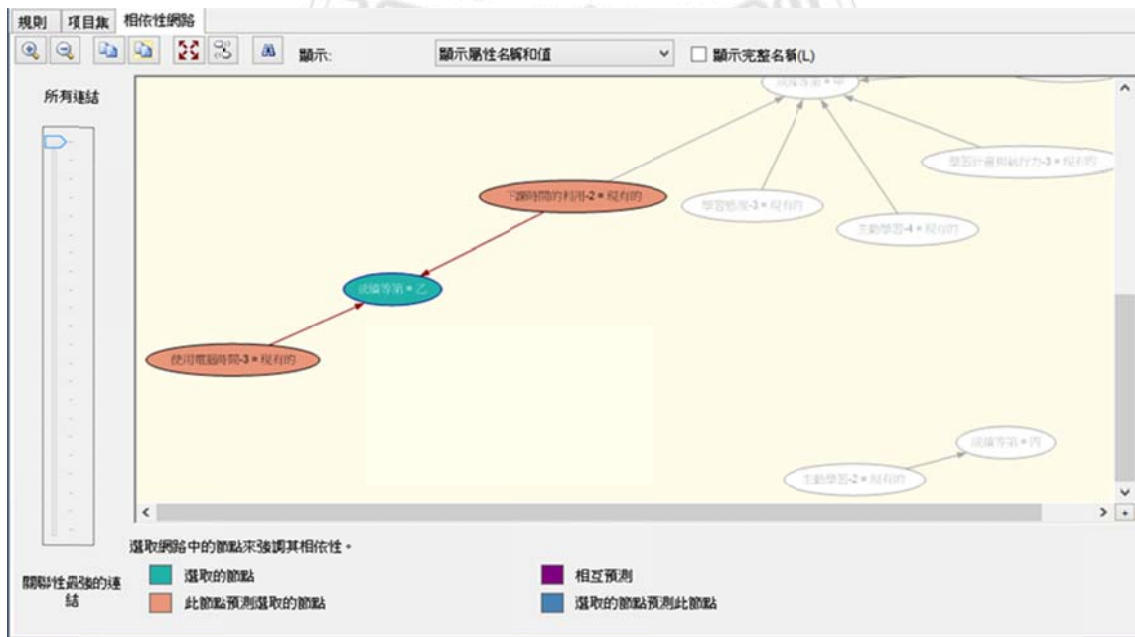


圖 4-21 「成績等第=乙」相依性網路圖

#### 四、「成績等第=丙」規則彙整

「篩選規則」輸入「成績等第=丙」，以圖 4-18 預設系統演算法參數執行採礦程序，符合規則有 1 條，符合研究所需，如附錄一的圖 4，整理規則如表 4-7，相依性網路如圖

4-22 所示。

表 4-7 「成績等第=丙」規則彙整表

機率	重要性	影響因素組合
0.316	0.551	主動學習-2

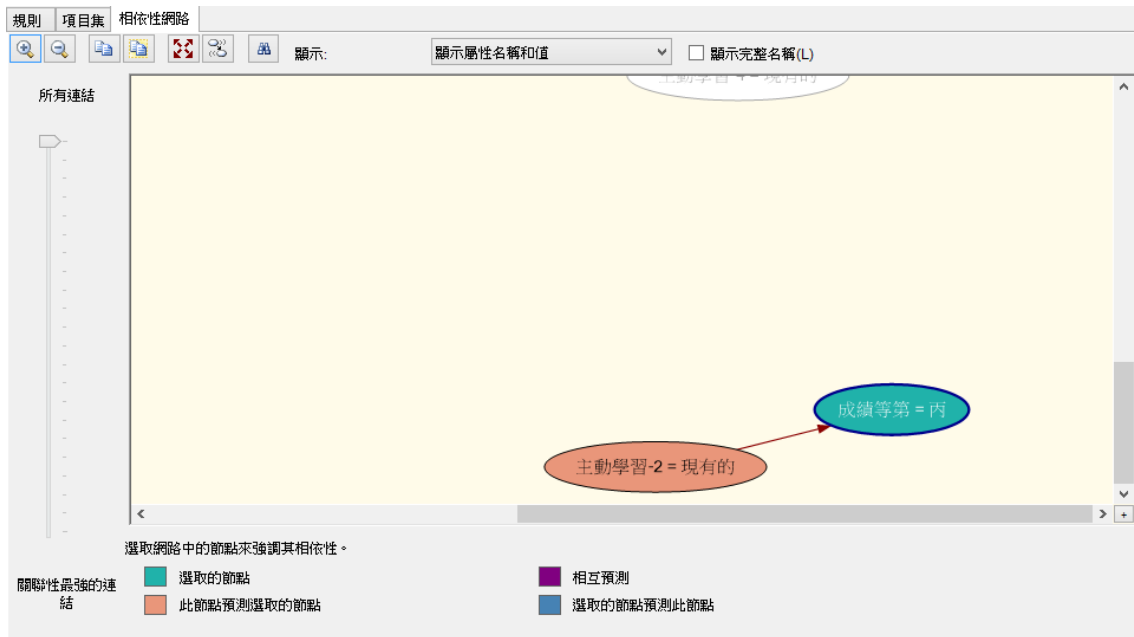


圖 4-22 「成績等第=丙」相依性網路圖

### 五、「成績等第=丁」規則彙整

「篩選規則」輸入「成績等第=丁」，以圖 4-18 預設統演算法參數執行採礦程序，無符合的規則。

最後，將附錄一的關聯規則進行彙整得到學生成績等第與個人學習背景因素關聯規則彙整表如表 4-8；表 4-8 中間的阿拉伯數字為背景因素變項值的選項編號與出現次數，例如：成績等第為甲的學生，其學習時間因素有選項 2(1 小時以上~未達 2 小時)也有選項 3(2 小時以上~未達 3 小時)的規則。

表 4-8 成績等第與個人學習背景因素關聯規則彙整表

學習背景因素	成績等第				
	優	甲	乙	丙	丁
1 學習時間	(3)*2	(2)*4 (3)			
2 休閒時間	(3)*1	(3)*1			
3 使用電腦時間	(2)*5	(1)*3	(3)*1		
4 下課時間的利用		(2)*3	(1)*1 (2)*4		
5 主動學習	(5)*10	(3)*1	(5)*2	(2)*1	
6 學習計畫與執行力	(1)*1 (4)*2	(3)*2	(1)*2 (3)*1		
7 作息規律性	(5)*3	(5)*2	(5)*2		
8 虛心求助頻率	(2)*3	(1)*7	(1)*1		
9 校外補習科目數	(4)*1	(3)*1 (4)*2	(1)*4		
10 學習態度	(4)*9	(4)*3	(3)*2 (4)*1		

## 貳、背景因素分析與探討

以下將分析與探討學生成績等第與影響成績的個人學習背景因素之間的關係：

### 一、學習時間

學習時間依成績等第(優、甲)順序，依序遞減 3(2 小時以上~未達 3 小時)、2(1 小時以上~未達 2 小時)，由此可觀察出學習時間對成績等第有正向的關聯。



## 二、 休閒時間

由「表 4-8」發現學生的休閒活動時間，似乎對成績等第沒有明顯的影響。

## 三、 使用電腦時間

學生要有好成績(甲等以上)，則平均每天使用電腦時間須為 1(未達 30 分鐘)或 2(30 分鐘以上~未達 1 小時)。當使用電腦時間為 3(1 小時以上~未達 2 小時)，則成績就較不理想。

## 四、 下課時間的利用

由「表 4-8」發現成績等第為甲、乙的學生下課時間的利用，大多為選項 2(20 分鐘以上~未達 30 分鐘)，但對成績等第沒有明顯的影響。

## 五、 主動學習

學生成績欲達優等，其主動學習的選項須為 5(3 天以上)，可見主動學習對學生成績是有影響的，讓學生養成一星期至少 3 天以上能主動學習的習慣對其得到好成績會有正面的幫助。

## 六、 學習計畫與執行力

學生安排學習計畫且依計畫執行的天數有隨成績等第(優、甲、乙)順序，依序遞減的趨勢，優 4(有安排學習計畫，平均一週有 3 天以上~未達 5 天會依計畫做功課、讀書)、甲 3(有安排學習計畫，平均一週有 1 天以上~未達 3 天會依計畫做功課、讀書)、乙 1(不曾安排學習計畫)，由此可

觀察出學習計畫與執行力和成績等第有關聯。

#### 七、 作息規律性

成績欲達中等(乙等)以上，其作息規律性的選項須為5(3天以上)，可見作息規律性對學生成績有正面的影響，當學生作息有一定規律時，學習時間與休息時間才能穩定。

#### 八、 虛心求助頻率

成績等第甲等以上的學生，其虛心求助頻率的選項須為1(平均10次中有8次以上會找人幫忙解決問題)或2(平均10次中有5次以上會找人幫忙解決問題)，可觀察出成績甲等以上的學生在學習上遇到問題時，虛心求助的機率是高於五成的。

#### 九、 校外補習科目數

成績等第欲達甲以上，則校外補習科目數選項須為3(2科)、或4(3科以上)，若未參加校外補習則成績就較不理想，可見校外補習科目數對成績有正面的影響。

#### 十、 學習態度

成績等第欲達甲以上，其學習態度選項須為4(3科以上)，可見認真學習的態度，對學生成績有正向的影響。

# 第五章 結論

## 第一節 研究結論

本論文以 Microsoft SQL Server 2012 資料採礦之關聯法則方法，探討國小四、五、六年級學生學期成績與學生個人的學習背景因素間的關聯性，已完成預定達成的目的。茲以整理研究所得到的關聯規則，提出以下結論：

- 一、學習時間對學生學期的成績有正向的關聯性，學生若要讓學期成績等第達甲以上，則建議其放學後，每日平均學習時間需要在 1 小時以上。
- 二、放學後平均每天使用電腦時間，在 1 小時以內，對於學期成績等第達到甲以上，具有正向的影響。
- 三、主動學習的習慣對學生的學期成績是有影響的，學期成績等第欲達到優，則須養成平均一星期裡面要有 3 天以上能主動學習的習慣。
- 四、學習計畫與執行力和學生的學期成績有關聯性，曾經安排學習計畫且平均一週有 3 天以上會依計畫做功課、讀書，對學生得到好成績有正面的幫助。
- 五、作息規律性對學生的學期成績有正向的影響，學生平均一星期裡面至少 3 天以上作息是有規律性的，對成績將有正面的幫助。
- 六、虛心求助對學生學期成績表現的影響是正向的，學生想要得到甲以上的學期成績等第，在學習遇到困難時，勇於發問與

虛心求助的平均機率需高於五成。

- 七、 校外補習的科目數對學生的成績也有正向的關聯性，參加校外補習或是請家教指導的科目數在 2 科以上，對學期成績會有正面的幫助。爰此，建議學校方面積極辦理補救教學、夜光天使等課後輔助計畫，以提昇弱勢家庭學生的學業成就。
- 八、 學習態度對學生的學期成績有正面的影響，學期成績等第欲達甲以上，則認真學習的科目需在 3 科以上。

## 第二節 後續研究建議

本論文研究對象僅限雲林縣某個案國小，研究結果概括性小，無法類推至不同地區的小學。如能從不同地區及不同類型國小抽樣做為研究樣本，則研究結果將更具價值性。

本論文僅對學生學期成績與個人的學習背景因素之間的相關性作探討，未來的研究可以針對各個學科與學生個人的學習背景因素之間的關係做探討，亦可以運用其它的資料採礦技術，例如：類神經網路、決策樹、貝氏機率等，來探討各因素之間的關係與關係之間的強弱，後續研究者可參考使用。

## 參考文獻

### 一、中文部份

- 王文科 (1991)。教育心理學。台北：五南。
- 王財印 (2000)。國民中學學生情緒智力、生活適應與學業成就關係之研究。國立高雄師範大學教育學系博士論文。
- 尹相志 (民 98)。SQL Server 2008 Data Mining。台北：悅知文化。
- 吳天泰 (1996)。特殊背景地區師資培育與進修。教育廳委託花蓮師範學院研究報告。
- 余民寧 (2006)。影響學習成就因素的探討。教育資料與研究雙月刊，73，11-24。
- 李美慧 (2004)。國小六年級學童個人背景、父母親管教方式與自我概念對學業成就之相關研究。長榮大學經營管理研究所碩士論文。
- 李義平 (2007)。影響國中生數學學習成績之因素分析—以台中市安和國中為例。中華大學應用數學系碩士論文。
- 周新富 (1999)。國中家庭背景、家庭文化資源、學校經驗與學習結果關係之研究。國立高雄師範大學教育系博士論文。
- 林建東 (2010)。應用資料採礦技術於國中學生的背景因素對在校成績影響之研究。南華大學資訊管理學系碩士論文。
- 郭生玉 (1981)。心理與教育研究法。台北：精華。
- 郭春悅 (2007)。家庭社經地位、同儕關係與學業成就之相關研究。國立屏東教育大學社會科教育學系碩士班碩士論文。
- 陳怡華 (2001)。國小學生家庭環境、閱讀動機與國語科學業成就之關係研究。國立高雄師範大學教育學系碩士論文。

詹德麟(民 97), 利用資料探勘技術於入學方式及學生背景對在校成績影響之研究。大葉大學事業經營研究所碩士論文。

鄭燕祥(2001)。學校效能及校本管理發展的機制。台北：心理。

## 二、西文部分

Gagn' e, R. M. (1985). The condition of learning theory of instruction. New York:Holt, Rinehert & Winston.

J. MacLennan, Z. H. Tang, and B. Crivat (2009), Data Mining with SQLServer® 2008, Wiley.



# 附錄一

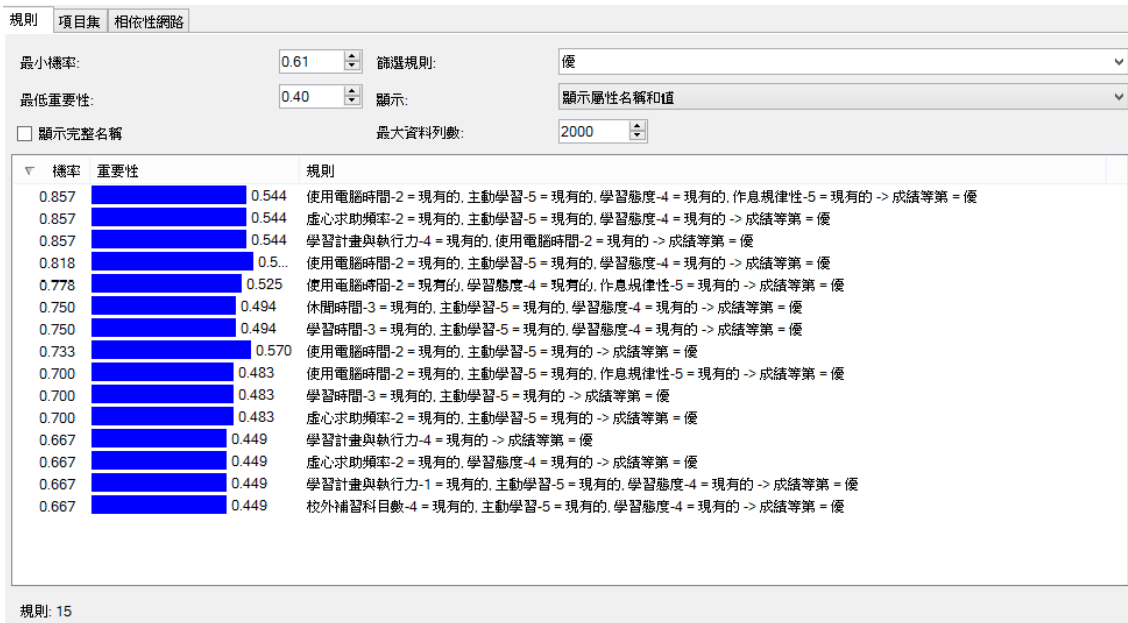


圖 1 「成績等第=優」規則

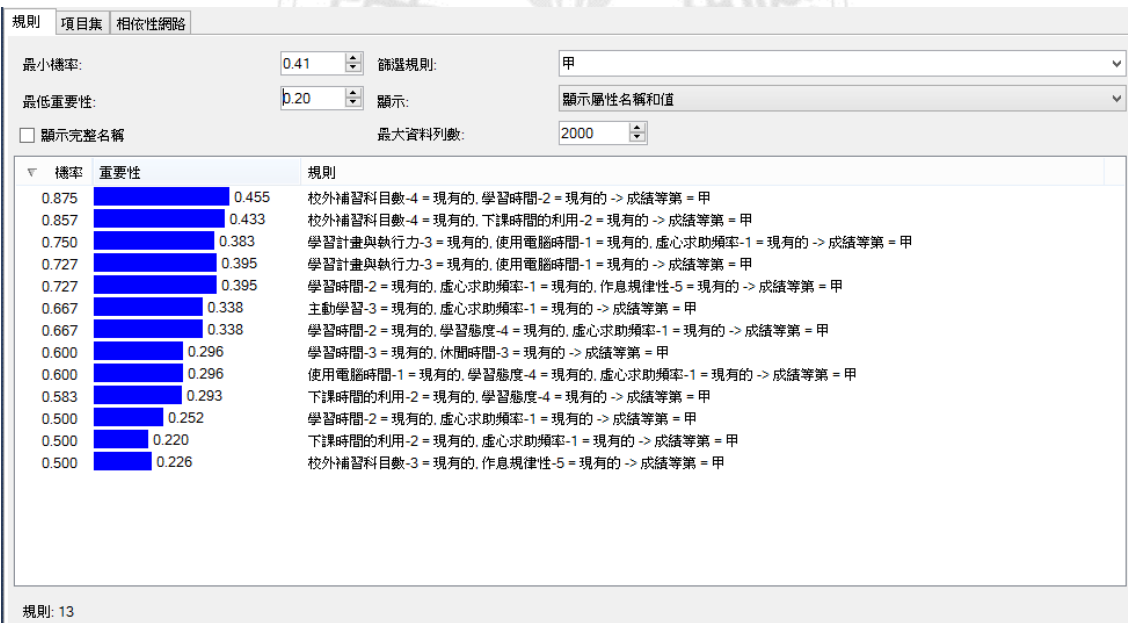


圖 2 「成績等第=甲」規則



圖 3 「成績等第=乙」規則

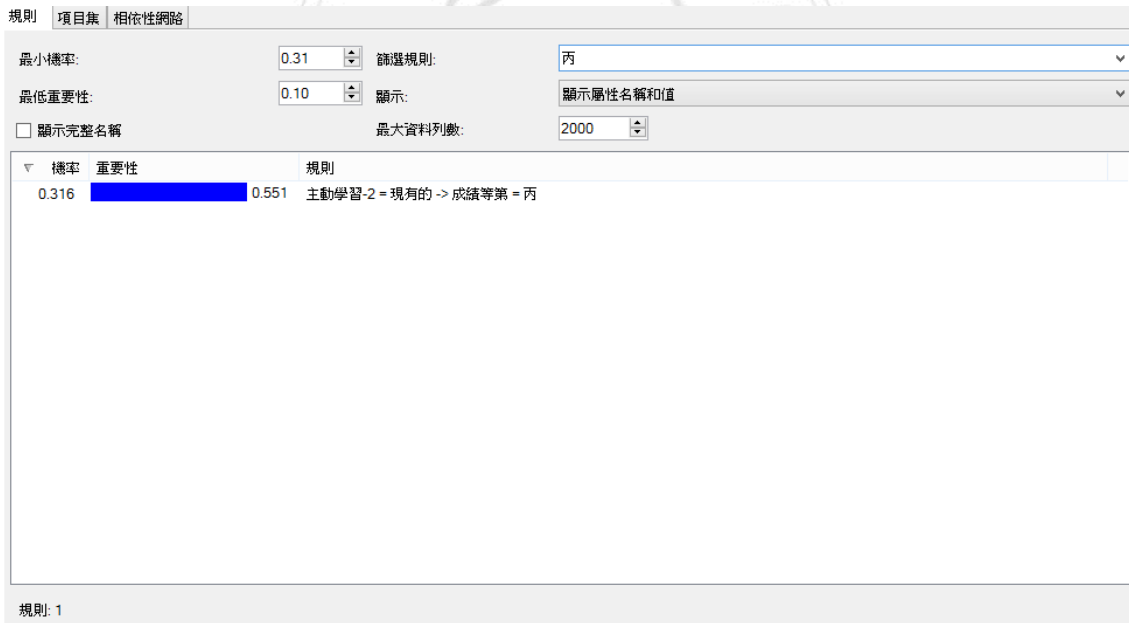


圖 4 「成績等第=丙」規則